Homotopy Algorithm for Optimal Control Problems with a Second-order State Constraint

Audrey Hermant

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Homotopy Algorithm for Optimal Control Problems with a Second-order State Constraint

Audrey Hermant

N° 6626

September 2008

Thème NUM

Rapport de recherche

ISSN 0249-6399 ISRN INRIA/RR--6626--FR+ENG

INSTITUT NATIONAL
DE RECHERCHE
EN INFORMATIQUE
ET EN AUTOMATIQUE

INRIA

centre de recherche
SACLAY – ÎLE-DE-FRANCE

Homotopy Algorithm for Optimal Control Problems with a Second-order State Constraint

Audrey Hermant*

Thème NUM — Systèmes numériques
Équipes-Projets Commands

Rapport de recherche n° 6626 — September 2008 — 31 pages

Abstract: This paper deals with optimal control problems with a regular second-order state constraint and a scalar control, satisfying the strengthened Legendre-Clebsch condition. We study the stability of structure of stationary points. It is shown that under a uniform strict complementarity assumption, boundary arcs are stable under sufficiently smooth perturbations of the data. On the contrary, nonreducible touch points are not stable under perturbations. We show that under some reasonable conditions, either a boundary arc or a second touch point may appear. Those results allow us to design an homotopy algorithm which automatically detects the structure of the trajectory and initializes the shooting parameters associated with boundary arcs and touch points.

Key-words: Optimal control, second-order state constraint, stability analysis, shooting algorithm, homotopy method.

* CMAP, École Polytechnique, INRIA Saclay Île-de-France, Route de Saclay, 91128 Palaiseau, France.

Centre de recherche INRIA Saclay – Île-de-France
Parc Orsay Université
4, rue Jacques Monod, 91893 ORSAY Cedex
Téléphone : +33 1 72 92 59 00

Algorithme d'homotopie pour les problèmes de commande optimale avec une contrainte sur l'état du second ordre

Résumé : Cet article est consacré aux problèmes de commande optimale avec une contrainte sur l'état scalaire du second ordre régulière et une commande scalaire, lorsque la condition forte de Legendre-Clebsch est satisfaite. On montre que sous une hypothèse de complémentarité stricte uniforme, les arcs frontières sont stables sous des perturbations suffisamment régulières des données. Au contraire, les points de contact isolés non réductibles ne sont pas stables. Sous des conditions raisonables, on montre que soit un arc frontière soit un second point de contact isolé peut apparaître. Ces résultats nous permettent de concevoir un algorithme d'homotopie qui détecte automatiquement la structure de la trajectoire et initialise les paramètres de tir associés aux arcs frontière et points de contact isolés.

Mots-clés : Commande optimale, contrainte sur l'état du second ordre, analyse de stabilité, algorithme de tir, méthode d'homotopie.

1 Introduction

This paper deals with optimal control problems with a state constraint of second-order (see [10, 22]). Many papers devoted to optimal control problems with state constraints deal with state constraints of first-order (see e.g. [15, 17, 18, 12, 19, 13, 7]), i.e. when the control appears explicitly after one time derivation of the state constraint along the dynamics. This assumption may not be satisfied in applications. For example, in the problem of the atmospheric reentry of a space shuttle, if the control is the bank angle (the angle of attack being fixed), the constraints on the thermal flux, normal acceleration and dynamic pressure are second-order state constraints, see [9].

When the strengthened Legendre-Clebsch condition holds, the shooting algorithm enables to solve optimal control problems with a very high accuracy at low cost. This algorithm (see [26]) is based on the parametrization of the trajectory by a finite-dimensional vector of *shooting parameters* and the resolution of the resulting multi-point boundary value problem by a Newton's method. Shooting methods are very sensitive to the initial conditions, and require a careful initialization of all parameters. Moreover, in presence of constraints, the structure of constraints (the number and order of boundary arcs and touch points) has to be known a priori. This makes the shooting algorithm generally hard to apply. However, when the precision is a strong requirement, such as e.g. to compute aerospace trajectories, shooting algorithms may be preferred to others methods, less accurate.

In order to determine the structure of the trajectory, which is generally unknown, and facilitate the initialization of parameters, homotopy (or continuation) methods can be used. Their well-known principle (see [1]) is to solve a sequence of problems depending continuously on a parameter, such that the first problem is "easy" to solve (e.g. the problem without the state constraint) and the last problem is the original problem. Doing so the structure of solutions may vary in the course of iterations. Homotopy methods have been applied to control problems with control constraints in e.g. [14, 21] and with state constraints in e.g. [4, 11]. The difficulty to apply classical continuation methods is connected with the changes of structure of the trajectory. Moreover, when the structure of the trajectory changes, the dimension of the vector of shooting parameters changes as well. In [7], an homotopy algorithm has been proposed for first-order state constraints, whose novelty is to automatically detect the changes in the structure of the trajectory and initialize the associated shooting parameters. It is well-known that the structure of a trajectory highly depends on the order of the constraint (see [10]). In this paper, we aim to extend the homotopy algorithm of [7] to second-order state constraints.

They are two main tools in the analysis of the homotopy method. Firstly, stability results which guarantee the existence and local uniqueness of a solution for the perturbed problem, and insure that the homotopy path is locally well-defined. Secondly, an analysis of the structure of solutions of the perturbed problem. New results concerning the first point (stability analysis) have been obtained recently in [16]. Contrary to previous stability results known for second- (and higher-)order state constraints ([20, 8]), no assumptions on the structure of the trajectory are made. This allows us precisely to deal with situations encountered in the homotopy method, when the structure of solution is not stable and hence, where the stability and sensitivity results of [20, 8] do not apply.

In this paper, results are obtained on the second point, i.e. we study the evolution of structure of solutions under small perturbations of the data. We show that when a strict complementarity hypothesis is satisfied on boundary arcs, then the latter are stable for a class of sufficiently smooth perturbations. Then we study the case of nonreducible touch points, which are excluded from the analysis based on shooting methods in [20] and [8]. In that case the structure of the trajectory is not stable. We show that under some rather general conditions, either a boundary arc or a second touch point may appear. Finally, we follows [7] in order to describe an homotopy method for second-order state constraints. The analysis is more involved than for first-order state constraints, since the structure of second-order state constraints is more complex (both essential touch points and boundary arcs are possible, while first-order state constraints typically do not have essential touch points).

The paper is organized as follows. Preliminaries (optimality conditions, assumptions) are recalled in section 2. In section 3, the stability of boundary arcs is studied. In section 4, the case of nonreducible touch points is dealt with. In section 5, the stability result of [16] is recalled. In section 6, lemmas used in the analysis of the homotopy method are given. In section 7, the homotopy algorithm is presented and analyzed. Finally, in section 8, some comments are given. The contributions of the paper are the structural analysis of stationary points in sections 3 and 4 and the analysis of the homotopy algorithm. The application of this homotopy algorithm to the atmospheric reentry of a space shuttle will be the subject of a forthcoming paper.

2 Preliminaries

We consider the following optimal control problem with a scalar control and scalar state constraint:

$$(\mathcal{P}) \qquad \min_{(u,y)\in\mathcal{U}\times\mathcal{Y}} \int_0^T \ell(u(t),y(t))\mathrm{d}t + \phi(y(T)) \tag{1}$$

$$\text{subject to} \qquad \dot{y}(t) \;=\; f(u(t),y(t)) \quad \text{for a.a. } t\in[0,T], \quad y(0)=y_0 \tag{2}$$

$$g(y(t)) \;\leq\; 0 \quad \text{for all } t\in[0,T] \tag{3}$$

with the control and state spaces $\mathcal{U} := L^\infty(0,T;\mathbb{R})$ and $\mathcal{Y} := W^{1,\infty}(0,T;\mathbb{R}^n)$. Throughout the paper, it is assumed that assumptions (A0) and (A1) below hold:

(A0) The data $\ell : \mathbb{R}\times\mathbb{R}^n \to \mathbb{R}$, $\phi : \mathbb{R}^n \to \mathbb{R}$, $f : \mathbb{R}\times\mathbb{R}^n \to \mathbb{R}^n$, (resp. $g : \mathbb{R}^n \to \mathbb{R}$) are C^3 (resp. C^4) mappings, with locally Lipschitz continuous third-order (resp. fourth-order) derivatives, and f is Lipschitz continuous.

(A1) The initial condition $y_0 \in \mathbb{R}^n$ satisfies $g(y_0) < 0$.

The state constraint is assumed to be of *second-order*. This means that the first-order time derivative $g^{(1)} : \mathbb{R}\times\mathbb{R}^n \to \mathbb{R}$ of the constraint, defined by

$$g^{(1)}(u,y) := g_y(y)f(u,y)$$

does not depend on the control variable u, i.e. $g_u^{(1)} \equiv 0$ (and hence, we may write $g^{(1)}(y) = g^{(1)}(u,y)$), and the second-order time derivative $g^{(2)} : \mathbb{R}\times\mathbb{R}^n \to \mathbb{R}$, defined by

$$g^{(2)}(u,y) := g_y^{(1)}(y)f(u,y)$$

depends explicitly on the control, i.e. $g_u^{(2)} \not\equiv 0$.

Notation We denote by subscripts Fréchet derivatives w.r.t. the variables u, y, i.e. $f_y(u,y) = D_y f(u,y)$, $f_{yy}(u,y) = D^2_{yy} f(u,y)$, etc. The derivative with respect to the time is denoted by a dot, i.e. $\dot{y} = \frac{\mathrm{d}y}{\mathrm{d}t} = y^{(1)}$. The set of row vectors of dimension n is denoted by $\mathbb{R}^{n*}$. Adjoint or transpose operators are denoted by the symbol $^\top$. The euclidean norm is denoted by $|\cdot|$. By $L^r(0,T)$ we denote the Lebesgue space of measurable functions such that $\|u\|_r := (\int_0^T |u(t)|^r \mathrm{d}t)^{1/r} < \infty$ for $1 \leq r < \infty$, $\|u\|_\infty := \mathrm{supess}_{[0,T]}\,|u(t)| < \infty$. The space $W^{s,r}(0,T)$ denotes the Sobolev space of functions in $L^r(0,T)$ having their s first weak derivatives in $L^r(0,T)$, and we denote by H^s the space $W^{s,2}$. The space of continuous functions over $[0,T]$ and its dual space, the space of bounded Borel measures, are denoted respectively by $C[0,T]$ and $\mathcal{M}[0,T]$. The cone of continuous functions with nonpositive values over $[0,T]$ is denoted by $K := C_-[0,T]$ and its dual space, the set of nonnegative measures, is denoted by $\mathcal{M}_+[0,T]$. The space of functions of bounded variation over $[0,T]$ is denoted by $BV[0,T]$, and the set of normalized BV functions vanishing at T is denoted by $BV_T[0,T]$. Functions of bounded variation are w.l.o.g. assumed to be right-continuous. We identify the elements of $\mathcal{M}[0,T]$ with the distributional derivatives $\mathrm{d}\eta$ of functions η in $BV_T[0,T]$. The support and the total variation of the measure $\mathrm{d}\eta \in \mathcal{M}[0,T]$ are denoted respectively by $\mathrm{supp}(\mathrm{d}\eta)$ and $|\mathrm{d}\eta|_{\mathcal{M}}$. Left- and right limits of a function of bounded variation φ will be denoted by $\varphi(\tau^\pm) := \lim_{t\to\tau^\pm}\varphi(t)$, and jumps by $[\varphi(\tau)] := \varphi(\tau^+) - \varphi(\tau^-)$. The cardinal of a finite set $\mathcal{T}$ is denoted by $|\mathcal{T}|$.

We call a *trajectory* an element $(u,y)\in\mathcal{U}\times\mathcal{Y}$ satisfying the state equation (2). A trajectory satisfying the state constraint (3) is said to be *feasible*. The contact set of a feasible trajectory is defined by

$$I(g(y)) := \{t\in[0,T] : g(y(t)) = 0\} \tag{4}$$

and for a small $\varepsilon > 0$, a neighborhood of the contact set is denoted by

$$I_\varepsilon(g(y)) := \{t \in [0,T] : \mathrm{dist}\{t, I(g(y))\} < \varepsilon\}. \tag{5}$$

A *boundary arc* (resp. *interior arc*) of a feasible trajectory (u,y) is a maximal (open) interval of positive measure (τ_1,τ_2) such that $g(y(t)) = 0$ (resp. $g(y(t)) < 0$) for all $t \in (\tau_1,\tau_2)$. The left- and right endpoints of a boundary arc (τ_{en},τ_{ex}) are called respectively *entry* and *exit* point. A *touch point* τ_{to} is an isolated contact point, i.e. such that $g(y(\tau_{to})) = 0$ and $g(y(t)) < 0$ for $t \neq \tau_{to}$ in a neighborhood of τ_{to}. An entry (resp. exit) point is said to be *regular*, if it belongs to $(0,T)$ and if there exists $\delta > 0$ such that $g(y(t)) < 0$ on $(\tau_{en} - \delta, \tau_{en})$ (resp. on $(\tau_{ex}, \tau_{ex} + \delta)$). A boundary arc is regular, if its entry and exit points are regular. The *structure* of a trajectory is the number and order of its boundary arcs and touch points.

Optimality conditions Let us first recall the well-known first-order necessary optimality condition of problem $(\mathcal{P})$. The *Hamiltonian* $H : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^{n*} \to \mathbb{R}$ is defined by

$$H(u,y,p) \;:=\; \ell(u,y) + pf(u,y). \tag{6}$$

We say that a feasible trajectory (u,y) is a *stationary point* of $(\mathcal{P})$, if there exists $(p,\eta) \in BV([0,T];\mathbb{R}^{n*}) \times BV_T[0,T]$ such that

$$\dot{y} \;=\; f(u,y), \qquad y(0) = y_0, \tag{7}$$
$$-\mathrm{d}p \;=\; H_y(u,y,p)\mathrm{d}t + g_y(y)\mathrm{d}\eta, \qquad p(T) = \phi_y(y(T)) \tag{8}$$
$$0 \;=\; H_u(u(t),y(t),p(t)) \qquad \text{a.e. on } [0,T] \tag{9}$$
$$0 \;\geq\; g(y(t)) \;\text{ for all } t \in [0,T], \quad \mathrm{d}\eta \in \mathcal{M}_+[0,T], \quad \mathrm{supp}(\mathrm{d}\eta) \subset I(g(y)). \tag{10}$$

Alternative formulation For the stability analysis, it is convenient to write the optimality condition using alternative multipliers η^2 and p^2, uniquely related to (p,η) in the following way:

$$\eta^1(t) \;:=\; \int_{(t,T]} \mathrm{d}\eta(s) = -\eta(t), \qquad \eta^2(t) := \int_t^T \eta^1(s)\mathrm{d}s, \tag{11}$$
$$p^2(t) \;:=\; p(t) - \eta^1(t)g_y(y(t)) - \eta^2(t)g_y^{(1)}(y(t)), \qquad t \in [0,T]. \tag{12}$$

We see that η^2 belongs to the set $BV_T^2[0,T]$, defined by

$$BV_T^2[0,T] := \{\xi \in W^{1,\infty}(0,T) : \xi(T) = 0,\ \dot{\xi} \in BV_T[0,T]\}. \tag{13}$$

Define the *alternative Hamiltonian* $\tilde{H} : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^{n*} \times \mathbb{R} \to \mathbb{R}$ by

$$\tilde{H}(u,y,p^2,\eta^2) \;:=\; H(u,y,p^2) + \eta^2 g^{(2)}(u,y), \tag{14}$$

where H is the classical Hamiltonian (6). Using these alternative multipliers, we obtain easily by a direct calculation (see e.g. [22] or [6, Lemma 3.4]) that a feasible trajectory $(u,y) \in \mathcal{U} \times \mathcal{Y}$ is a stationary point of $(\mathcal{P})$ iff there exists $(p^2,\eta^2) \in W^{1,\infty}(0,T;\mathbb{R}^{n*}) \times BV_T^2[0,T]$ such that

$$\dot{y}(t) \;=\; f(u(t),y(t)) \quad \text{a.e. on } [0,T], \quad y(0) = y_0, \tag{15}$$
$$-\dot{p}^2(t) \;=\; \tilde{H}_y(u(t),y(t),p^2(t),\eta^2(t)) \quad \text{a.e. on } [0,T], \quad p^2(T) = \phi_y(y(T)) \tag{16}$$
$$0 \;=\; \tilde{H}_u(u(t),y(t),p^2(t),\eta^2(t)) \quad \text{a.e. on } [0,T] \tag{17}$$
$$0 \;\geq\; g(y(t)) \;\text{ for all } t \in [0,T], \quad \mathrm{d}\dot{\eta}^2 \in \mathcal{M}_+[0,T], \quad \mathrm{supp}(\mathrm{d}\dot{\eta}^2) \subset I(g(y)). \tag{18}$$

Assumptions Let $(\bar{u},\bar{y})$ be a stationary point of $(\mathcal{P})$, with alternative multipliers $(\bar{p}^2,\bar{\eta}^2)$. We make the following assumptions:

(A2) The state constraint is a regular second-order state constraint, i.e. $g_u^{(1)} \equiv 0$ and

$$\exists\, \beta, \sigma > 0, \quad |g_u^{(2)}(\bar{u}(t), \bar{y}(t))| \geq \beta \quad \text{for a.a. } t \in I_\sigma(g(\bar{y})). \tag{19}$$

(A3) $\bar{u}$ is continuous on $[0, T]$ and the strengthened Legendre-Clebsch condition holds:

$$\exists\, \alpha > 0, \quad \tilde{H}_{uu}(\bar{u}(t), \bar{y}(t), \bar{p}^2(t), \bar{\eta}^2(t)) \geq \alpha \quad \text{for all } t \in [0, T]. \tag{20}$$

Lemma 2.1. *Let $(\bar{u}, \bar{y})$ be a stationary point of $(\mathcal{P})$ with alternative multipliers $(\bar{p}^2, \bar{\eta}^2)$ satisfying (A2)–(A3). Then $\bar{u}$ and $\bar{\eta}^2$ are of class C^2 on the interior of the (interior and boundary) arcs of the trajectory, with Lipschitz continuous second-order time derivatives.*

Proof. By the implicit function theorem applied to (17) on interior arcs, using that $\dot{\eta}^2$ is constant, and to $g^{(2)}(u(t), y(t)) = 0$ and (17) on boundary arcs, the control and alternative state constraint multipliers can be expressed, on the interior of arcs, as C^2 functions of the state and alternative costate (y, p^2). The result follows. □

Assume now that $(\bar{u}, \bar{y})$ has a (regular) boundary arc $(\bar{\tau}_{en}, \bar{\tau}_{ex})$. We consider the uniform strict complementarity assumption on boundary arcs below:

$$\exists\, \beta > 0, \qquad \ddot{\bar{\eta}}^2(t) \geq \beta \qquad \text{on } (\bar{\tau}_{en}, \bar{\tau}_{ex}). \tag{21}$$

Remark 2.2. Using the classical multipliers $(\bar{p}, \bar{\eta})$ associated with $(\bar{u}, \bar{y})$ in (7)–(10), assumption (21) can equivalently be rewritten as (recall that $\bar{\eta} = \dot{\bar{\eta}}^2$):

$$\exists\, \beta > 0, \qquad \frac{\mathrm{d}\bar{\eta}}{\mathrm{d}t}(t) \geq \beta \qquad \text{on } (\bar{\tau}_{en}, \bar{\tau}_{ex}). \tag{22}$$

Lemma 2.3. *Let $(\bar{u}, \bar{y})$ be a stationary point of $(\mathcal{P})$ satisfying (A2)–(A3) and having a regular boundary arc $(\bar{\tau}_{en}, \bar{\tau}_{ex})$. Then the uniform strict complementarity assumption* (21) *implies that*

$$\frac{\mathrm{d}^3}{\mathrm{d}t^3} g(\bar{y}(t))|_{t=\bar{\tau}_{en}^-} > 0, \qquad \frac{\mathrm{d}^3}{\mathrm{d}t^3} g(\bar{y}(t))|_{t=\bar{\tau}_{ex}^+} < 0. \tag{23}$$

For convenience, Lemma 2.3 will be proved in section 3, after the suitable notation has been introduced.

Perturbed optimal control problem We consider perturbed problems in the following form:

$$(\mathcal{P}^\mu) \qquad \min_{(u,y) \in \mathcal{U} \times \mathcal{Y}} \int_0^T \ell^\mu(u(t), y(t))\mathrm{d}t + \phi^\mu(y(T)) \tag{24}$$

$$\text{subject to} \qquad \dot{y}(t) = f^\mu(u(t), y(t)) \quad \text{a.e. on } [0, T], \quad y(0) = y_0^\mu \tag{25}$$

$$g^\mu(y(t)) \leq 0 \quad \text{for all } t \in [0, T]. \tag{26}$$

Here μ is the perturbation parameter, living in an open subset M_0 of a Banach space M. In what follows, we consider *stable extensions* $(\mathcal{P}^\mu)$ of problem $(\mathcal{P})$ in the following sense.

Definition 2.4. We say that $(\mathcal{P}^\mu)$ is a *stable extension* of $(\mathcal{P})$ if:
(i) There exists $\bar{\mu} \in M_0$ such that $(\mathcal{P}^{\bar{\mu}}) \equiv (\mathcal{P})$;
(ii) The mappings $\mathbb{R} \times \mathbb{R}^n \times M_0 \to \mathbb{R}$, $(u, y, \mu) \mapsto \ell^\mu(u, y)$; $\mathbb{R}^n \times M_0 \to \mathbb{R}$, $(y, \mu) \mapsto \phi^\mu(y)$; $M_0 \to \mathbb{R}^n$, $\mu \mapsto y_0^\mu$; $\mathbb{R} \times \mathbb{R}^n \times M_0 \to \mathbb{R}^n$, $(u, y, \mu) \mapsto f^\mu(u, y)$ (resp. $\mathbb{R}^n \times M_0 \to \mathbb{R}$, $(y, \mu) \mapsto g^\mu(y)$) are of class C^3 (resp. C^4), with locally Lipschitz continuous third-order (resp. fourth-order) derivatives, uniformly w.r.t. $\mu \in M_0$;
(iii) The dynamics f^μ is uniformly Lipschitz continuous over $\mathbb{R} \times \mathbb{R}^n$ for all $\mu \in M_0$;
(iv) The state constraint is not of first-order, i.e. $(g^\mu)_u^{(1)}(u, y) \equiv 0$ for all $(u, y, \mu) \in \mathbb{R} \times \mathbb{R}^n \times M_0$.

Abstract formulation Given a stable extension $(\mathcal{P}^\mu)$, the mapping $\mathcal{U} \times M_0 \to \mathcal{Y}$, $(u, \mu) \mapsto y_u^\mu$, where y_u^μ is the unique solution in $\mathcal{Y}$ of the state equation (25), is well-defined, and we may write the following abstract formulation of $(\mathcal{P}^\mu)$

$$\min_{u \in \mathcal{U}} J^\mu(u), \qquad G^\mu(u) \in K, \tag{27}$$

with the cost function $J^\mu(u) := \int_0^T \ell^\mu(u, y_u^\mu)\mathrm{d}t + \phi^\mu(y_u^\mu(T))$, the constraint mapping $G^\mu(u) := g^\mu(y_u^\mu)$, and the constraint cone $K = C_-[0,T]$.

Given a stationary point $(\bar{u}, \bar{y})$ of $(\mathcal{P})$, we say that the *uniform quadratic growth* condition holds, if for all stable extensions $(\mathcal{P}^\mu)$ of $(\mathcal{P})$, there exists $c, \rho > 0$ and a neighborhood $\mathcal{N}$ of $\bar{\mu}$, such that for any stationary point (u^μ, y^μ) of $(\mathcal{P}^\mu)$ with $\mu \in \mathcal{N}$ and $\|u^\mu - \bar{u}\|_\infty < \rho$,

$$J^\mu(u) \geq J^\mu(u^\mu) + c\|u - u^\mu\|_2^2, \quad \text{for all } u \in \mathcal{U} : G^\mu(u) \in K,\ \|u - \bar{u}\|_\infty < \rho. \tag{28}$$

Qualification condition and stability of multipliers Robinson's constraint qualification for problem $(\mathcal{P})$ in abstract form (27) is as follows (omitting the perturbation parameter at the reference point $\mu = \bar{\mu}$):

$$\exists\, \varepsilon > 0, \qquad \varepsilon B_{C[0,T]} \subset G(\bar{u}) + DG(\bar{u})\mathcal{U} - K, \tag{29}$$

where $B_{C[0,T]}$ denotes the open unit ball of the space $C[0,T]$. It is well-known that a local solution (weak minimum) of $(\mathcal{P})$ satisfying (29) is a stationary point of $(\mathcal{P})$.

Given $v \in L^r(0,T)$, $1 \leq r \leq \infty$, denote by z_v the unique solution in $W^{1,r}(0,T;\mathbb{R}^n)$ of the linearized state equation

$$\dot{z}_v(t) = f_y(\bar{u}(t), \bar{y}(t))z_v(t) + f_u(\bar{u}(t), \bar{y}(t))v(t) \quad \text{ a.e. on } [0,T], \quad z_v(0) = 0. \tag{30}$$

Assumption (A2) implies that Robinson's constraint qualification (29) holds, and that the multipliers associated with a stationary point are unique. This is a consequence of the lemma below.

Lemma 2.5 ([5, Prop. 10])**.** *Let $(\bar{u}, \bar{y})$ be a feasible trajectory of $(\mathcal{P})$ satisfying (A2). Then for all $r \in [1, +\infty]$ and all $\varepsilon \in (0, \sigma)$, with the σ of (19), so small that $\Omega_\varepsilon \subset [a, T]$ for some $a > 0$, the linear mapping*

$$L^r(0,T) \to W^{2,r}(\Omega_\varepsilon), \qquad v \mapsto (g_y(\bar{y}(\cdot))z_v(\cdot))|_{\Omega_\varepsilon}, \tag{31}$$

where $|_{\Omega_\varepsilon}$ denotes the restriction to the set Ω_ε, is onto, and therefore has a bounded right inverse by the open mapping theorem.

Let us end this section by recalling two results that will be used in the paper.

Proposition 2.6 ([16, Prop. 4.4])**.** *Let $(\bar{u}, \bar{y})$ be a stationary point of $(\mathcal{P})$ satisfying (A2). Then for every stable extension $(\mathcal{P}^\mu)$ of $(\mathcal{P})$ and for every stationary point (u, y) of $(\mathcal{P}^\mu)$, with (unique) associated multipliers (p, η) and alternative multipliers (p^2, η^2) given by (11)–(12), we have:*
(i) *If $\|\mu - \bar{\mu}\|, \|u - \bar{u}\|_\infty$,are small enough, then $\mathrm{d}\eta$ is uniformly bounded in $\mathcal{M}[0,T]$;*
Moreover, when $\|\mu - \bar{\mu}\|, \|u - \bar{u}\|_\infty \to 0$:
(ii) *$\mathrm{d}\eta$ weakly-* converges to $\mathrm{d}\bar{\eta}$ in $\mathcal{M}[0,T]$;*
(iii) *p^2 and η^2 converge uniformly to $\bar{p}^2$ and $\bar{\eta}^2$, respectively.*

Given $A, B \subset [0,T]$, we denote by $\mathrm{exc}\{A, B\}$ the *Hausdorff excess* of A over B, defined by

$$\mathrm{exc}\{A, B\} \ := \ \sup_{t \in A} \inf_{s \in B} |t - s|, \tag{32}$$

with the convention $\mathrm{exc}\{\emptyset, B\} = 0$.

Lemma 2.7 ([16, Lemma 4.6])**.** *Let $\mathrm{d}\bar{\eta} \in \mathcal{M}[0,T]$, and a sequence $(\mathrm{d}\eta_n) \subset \mathcal{M}[0,T]$ be such that $\mathrm{d}\eta_n$ weakly-* converges to $\mathrm{d}\bar{\eta}$ in $\mathcal{M}[0,T]$. Then $e_n := \mathrm{exc}\{\mathrm{supp}(\mathrm{d}\bar{\eta}), \mathrm{supp}(\mathrm{d}\eta_n)\}$ converges to zero when $n \to +\infty$.*

3 Stability of boundary arcs

The aim of this section is to show that boundary arcs are "stable" under perturbations, for sufficiently smooth perturbations (the *stable extensions* satisfying Def. 2.4). Here is the main result of this section.

Theorem 3.1. *Let $(\bar{u}, \bar{y})$ be a stationary point of $(\mathcal{P})$ satisfying (A2)–(A3). Assume that $(\bar{u}, \bar{y})$ has a regular boundary arc $(\bar{\tau}_{en}, \bar{\tau}_{ex})$ and that (21) holds. Then, for every stable extension $(\mathcal{P}^\mu)$ of $(\mathcal{P})$ and for all small enough $\delta > 0$, there exist $\rho, \varrho > 0$ such that if (u, y) is a stationary point of $(\mathcal{P}^\mu)$ with $\|\mu - \bar{\mu}\| < \varrho$ and $\|u - \bar{u}\|_\infty < \rho$, then (u, y) has on $(\bar{\tau}_{en} - \delta, \bar{\tau}_{ex} + \delta)$ a unique boundary arc (τ_{en}, τ_{ex}) (and no touch point). Moreover, we have that $|\tau_{en} - \bar{\tau}_{en}|, |\tau_{ex} - \bar{\tau}_{ex}| < \delta$ and (u, y) satisfies the uniform strict complementarity assumption* (21) *on (τ_{en}, τ_{ex}).*

We derive next some useful relations for the proof of Th. 3.1 and Lemma 2.3, and for other results of the paper. Let $(\bar{u}, \bar{y})$ be a stationary point of $(\mathcal{P}) \equiv (\mathcal{P}^{\bar{\mu}})$ satisfying (A2)–(A3) with alternative multipliers $(\bar{p}^2, \bar{\eta}^2)$, and let (u, y) be a stationary point of $(\mathcal{P}^\mu)$ with alternative multipliers (p^2, η^2). If $\|\mu - \bar{\mu}\|$ and $\|u - \bar{u}\|_\infty$ are small enough, then by (19), (20), and Prop. 2.6(iii), we have that

$$|(g^\mu)_u^{(2)}(u, y)| \;\geq\; \beta/2 \;>\; 0, \qquad \text{a.e. on } I_\sigma(g(\bar{y})) \supset I(g^\mu(y)), \tag{33}$$
$$\tilde{H}^\mu_{uu}(u, y, p^2, \eta^2) \;\geq\; \alpha/2 \;>\; 0 \qquad \text{on } [0, T], \tag{34}$$

with $\tilde{H}^\mu$ the alternative Hamiltonian (14) for $(\mathcal{P}^\mu)$. Moreover, by the implicit function theorem applied locally to (17) under hypothesis (A3), we may write that $u(t) = \Upsilon(y(t), p^2(t), \eta^2(t))$ for some C^2 function Υ, and hence u is continuous over $[0, T]$. It follows from Lemma 2.1 that u and η^2 are C^2 on the interior of arcs of the trajectory (u, y). So we may consider the time derivatives of the state constraint of order 3 and 4, defined on the interior of (interior and boundary) arcs by:

$$(g^\mu)^{(3)}(\dot{u}, u, y) \;:=\; (g^\mu)_u^{(2)}(u, y)\dot{u} + (g^\mu)_y^{(2)}(u, y) f^\mu(u, y) \tag{35}$$
$$(g^\mu)^{(4)}(\ddot{u}, \dot{u}, u, y) \;:=\; (g^\mu)_u^{(2)}(u, y)\ddot{u} + (g^\mu)_u^{(3)}(\dot{u}, u, y)\dot{u} + (g^\mu)_y^{(3)}(\dot{u}, u, y) f^\mu(u, y). \tag{36}$$

Time derivations of (17) shows that, on the interior of arcs, where u and η^2 are C^2 (arguments (u, y, p^2, η^2) and time are omitted as well as the superscript μ to simplify the notation)

$$0 \;=\; \tilde{H}_{uu}\dot{u} + \tilde{H}_{uy} f - \tilde{H}_y f_u + \dot{\eta}^2 g_u^{(2)} \tag{37}$$
$$0 \;=\; \tilde{H}_{uu}\ddot{u} + \ddot{\eta}^2 g_u^{(2)} + \Phi_1(\dot{u}, \dot{\eta}^2, u, y, p^2, \eta^2, \mu), \tag{38}$$

where Φ_1 is a locally Lipschitz continuous function w.r.t. its arguments. By (34), multiplying (38) by $g_u^{(2)}/\tilde{H}_{uu}$ and using (36) we may write that for all $t \in (0, T)$ in the interior of arcs,

$$0 \;=\; g^{(4)} + \frac{(g_u^{(2)})^2}{\tilde{H}_{uu}}\ddot{\eta}^2 + \Phi_2(\dot{u}, \dot{\eta}^2, u, y, p^2, \eta^2, \mu), \tag{39}$$

where Φ_2 is a locally Lipschitz continuous function w.r.t. its arguments. Moreover, by (33), it follows from (35) and (37) that we may express $\dot{u}$ and $\dot{\eta}^2$ as locally Lipschitz continuous functions of $(g^{(3)}, u, y, p^2, \eta^2, \mu)$, i.e. more precisely

$$\dot{u} \;=\; (g_u^{(2)})^{-1}(g^{(3)} - g_y^{(2)} f),$$
$$\dot{\eta}^2 \;=\; -(g_u^{(2)})^{-1}(\tilde{H}_{uu}(g_u^{(2)})^{-1}(g^{(3)} - g_y^{(2)} f) + \tilde{H}_{uy} f - \tilde{H}_y f_u).$$

Therefore, (39) yields, on the interior of arcs,

$$g^{(4)} + \frac{(g_u^{(2)})^2}{\tilde{H}_{uu}}\ddot{\eta}^2 + \Lambda(g^{(3)}, u, y, p^2, \eta^2, \mu) \;=\; 0 \tag{40}$$

where Λ is a locally Lipschitz continuous function w.r.t. its arguments.

In the sequel, we abbreviate the notation as follows:

$$\begin{aligned}
g^{(3)}(t) &:= (g^{\mu})^{(3)}(\dot u(t), u(t), y(t)), \qquad g^{(4)}(t) := (g^{\mu})^{(4)}(\ddot u(t), \dot u(t), u(t), y(t)) && (41)\\
\bar g^{(3)}(t) &:= (g^{\bar\mu})^{(3)}(\dot{\bar u}(t), \bar u(t), \bar y(t)), \qquad \bar g^{(4)}(t) := (g^{\bar\mu})^{(4)}(\ddot{\bar u}(t), \dot{\bar u}(t), \bar u(t), \bar y(t)), && (42)\\
\tilde H_{uu}(t) &:= \tilde H^{\mu}_{uu}(u(t), y(t), p^2(t), \eta^2(t)), \quad \bar H_{uu}(t) := \tilde H^{\bar\mu}_{uu}(\bar u(t), \bar y(t), \bar p^2(t), \bar\eta^2(t)), && (43)\\
\Lambda(t) &:= \Lambda(g^{(3)}(t), u(t), y(t), p^2(t), \eta^2(t), \mu),\\
\bar\Lambda(t) &:= \Lambda(\bar g^{(3)}(t), \bar u(t), \bar y(t), \bar p^2(t), \bar\eta^2(t), \bar\mu).
\end{aligned}$$

We start by the proof of Lemma 2.3 and then give that of Th. 3.1.

Proof of Lemma 2.3. Assume that (21) holds. Assume by contradiction that (23) does not hold, i.e. $\bar g^{(3)}$ is continuous at entry or exit point τ. Then by continuity of $(\bar u, \bar y, \bar p^2, \bar\eta^2)$, (40) implies that

$$[\bar g^{(4)}(\tau)] + \frac{(\bar g_u^{(2)})^2}{\bar H_{uu}}[\ddot{\bar\eta}^2(\tau)] = 0. \tag{44}$$

Since $\bar g^{(3)}$ is continuous at τ, in the neighborhood of τ, on the side of the interior arc, we have

$$g(\bar y(t)) = \bar g^{(4)}(\tau^{\pm})\frac{(t-\tau)^4}{24} + o((t-\tau)^4) \ \le\ 0,$$

where $\tau^{\pm}$ denotes τ^- if $\tau = \bar\tau_{en}$ and τ^+ if $\tau = \bar\tau_{ex}$. Since $\bar g^{(4)} = 0$ on the interior of the boundary arc, it follows that

$$[\bar g^{(4)}(\bar\tau_{en})] \ge 0 \qquad \text{and} \qquad [\bar g^{(4)}(\bar\tau_{ex})] \le 0. \tag{45}$$

Moreover, (21) implies that

$$[\ddot{\bar\eta}^2(\bar\tau_{en})] \ge \beta > 0 \qquad \text{and} \qquad [\ddot{\bar\eta}^2(\bar\tau_{ex})] \le -\beta < 0. \tag{46}$$

Since $\frac{(\bar g_u^{(2)})^2}{\bar H_{uu}} > 0$ by (A2)–(A3), the above display and (44) yield

$$[\bar g^{(4)}(\bar\tau_{en})] < 0 \qquad \text{and} \qquad [\bar g^{(4)}(\bar\tau_{ex})] > 0,$$

contradicting (45). Therefore (23) holds, which completes the proof. □

Proof of Th. 3.1. Let (u, y) be a stationary point of $(\mathcal{P}^{\mu})$ with u in a L^∞-neighborhood of $\bar u$ and μ in a neighborhood of $\bar\mu$. Assume by contradiction that (u, y) has an interior arc $(\tau_1, \tau_2) \subset (\bar\tau_{en} - \delta, \bar\tau_{ex} + \delta)$. On the interior arc (τ_1, τ_2), u and η^2 are C^2, and $g(t) := g^{\mu}(y(t))$ attains its minimum on (τ_1, τ_2) at a point where the second-order derivative $g^{(2)}$ is nonnegative. Since $g^{(2)}(\tau_i) \le 0$, $i = 1, 2$, the continuous function $g^{(2)}$ attains its maximum over $[\tau_1, \tau_2]$ at some point $t_m \in (\tau_1, \tau_2)$, and we have at this point of maximum of $g^{(2)}$

$$g^{(3)}(t_m) \ =\ 0 \qquad \text{and} \qquad g^{(4)}(t_m) \ \le\ 0. \tag{47}$$

Assume first that $t_m \in (\bar\tau_{en}, \bar\tau_{ex})$. By Prop. 2.6(iii), $(y, p^2, \eta^2) \to (\bar y, \bar p^2, \bar\eta^2)$ uniformly over $[0, T]$ when $\|\mu - \bar\mu\| \to 0$ and $\|u - \bar u\|_\infty \to 0$, and $g^{(3)}(t_m) = 0 = \bar g^{(3)}(t_m)$ since $t_m \in (\bar\tau_{en}, \bar\tau_{ex})$. Therefore, $\Lambda(t_m) - \bar\Lambda(t_m) \to 0$, and hence (40) implies that when $\|\mu - \bar\mu\| \to 0$ and $\|u - \bar u\|_\infty \to 0$,

$$g^{(4)}(t_m) + \frac{(g_u^{(2)})^2}{\tilde H_{uu}}\ddot\eta^2(t_m) \ - \ (\bar g^{(4)}(t_m) + \frac{(\bar g_u^{(2)})^2}{\bar H_{uu}}\ddot{\bar\eta}^2(t_m)) \ \to\ 0.$$

But $\ddot\eta^2(t_m) = 0$ since we are on an interior arc for (u, y), and $\bar g^{(4)}(t_m) = 0$ since we are on a boundary arc for $(\bar u, \bar y)$. It follows that when $\|\mu - \bar\mu\| \to 0$ and $\|u - \bar u\|_\infty \to 0$,

$$g^{(4)}(t_m) \ - \ \frac{(\bar g_u^{(2)})^2}{\bar H_{uu}}\ddot{\bar\eta}^2(t_m) \ \to\ 0.$$

Since $\frac{(\bar{g}_u^{(2)})^2}{\bar{H}_{uu}} \geq C > 0$ by (19) and (20), we obtain by (21) that $\frac{(\bar{g}_u^{(2)})^2}{\bar{H}_{uu}}\ddot{\bar{\eta}}^2(t_m) \geq C\beta > 0$. Therefore, for $\|\mu - \bar{\mu}\|$ and $\|u - \bar{u}\|_\infty$ small enough, $g^{(4)}(t_m) \geq C\beta/2 > 0$, contradicting (47).

Assume now that $t_m \in (\bar{\tau}_{en} - \delta, \bar{\tau}_{en}]$ (the case when $t_m \in [\bar{\tau}_{ex}, \bar{\tau}_{ex} + \delta)$ is analogous). For all $0 < \varepsilon < \delta$, if $\|\mu - \bar{\mu}\|$ and $\|u - \bar{u}\|_\infty$ are small enough, then $g^\mu(y(t)) < 0$ on the interval $[\bar{\tau}_{en} - \delta, \bar{\tau}_{en} - \varepsilon]$. This implies that $t_m \uparrow \bar{\tau}_{en}$ when $\|\mu - \bar{\mu}\| \to 0$ and $\|u - \bar{u}\|_\infty \to 0$. Therefore, since $g^{(3)}(t_m) = 0 = \bar{g}^{(3)}(\bar{\tau}_{en}^+)$ and $(\bar{u}, \bar{y}, \bar{p}^2, \bar{\eta}^2)$ is continuous over $[0, T]$, we obtain by Prop. 2.6(iii) that $\Lambda(t_m) \to \bar{\Lambda}(\bar{\tau}_{en}^+)$. It follows then from (40) that

$$g^{(4)}(t_m) + \frac{(g_u^{(2)})^2}{\tilde{H}_{uu}}\ddot{\eta}^2(t_m) \ \to \ \bar{g}^{(4)}(\bar{\tau}_{en}^+) + \frac{(\bar{g}_u^{(2)})^2}{\bar{H}_{uu}}\ddot{\bar{\eta}}^2(\bar{\tau}_{en}^+) \ \geq \ 0 + C\beta > 0,$$

contradicting (47) again since $g^{(4)}(t_m) \leq 0$ and $\ddot{\eta}^2(t_m) = 0$. This shows that for all small $\delta > 0$, if $\|u - \bar{u}\|_\infty$ and $\|\mu - \bar{\mu}\|$ are small enough, then (u, y) has *no interior arc* contained in $(\bar{\tau}_{en} - \delta, \bar{\tau}_{ex} + \delta)$.

It follows that $I(g^\mu(y)) \cap (\bar{\tau}_{en} - \delta, \bar{\tau}_{ex} + \delta)$ is either empty, or a touch point, or a boundary arc. Let us refute the two first possibilities. For all small $\varepsilon > 0$, if $\|u - \bar{u}\|_\infty$ and $\|\mu - \bar{\mu}\|$ are small enough, then $I(g^\mu(y)) \subset I_\varepsilon(g(\bar{y}))$. By hypothesis (21), Prop. 2.6(ii) and Lemma 2.7 (recall that $\mathrm{d}\dot{\eta}^2 = \mathrm{d}\eta$), for all $t \in [\bar{\tau}_{en}, \bar{\tau}_{ex}] \subset \mathrm{supp}(\mathrm{d}\bar{\eta})$, there exists $s \in \mathrm{supp}(\mathrm{d}\dot{\eta}^2) \subset I(g^\mu(y))$ such that $|t - s| < \varepsilon$. Therefore, we deduce that $I(g^\mu(y)) \cap (\bar{\tau}_{en} - \delta, \bar{\tau}_{ex} + \delta)$ is necessarily a boundary arc (τ_{en}, τ_{ex}), and that $|\tau_{en} - \bar{\tau}_{en}|, |\tau_{ex} - \bar{\tau}_{ex}| < \varepsilon$.

It remains to show that uniform strict complementarity holds on that boundary arc. By (40), it holds for all t in boundary arc (τ_{en}, τ_{ex}) that

$$\ddot{\eta}^2(t) = -\frac{\tilde{H}_{uu}(t)}{(g_u^{(2)}(t))^2}\Lambda(0, u(t), y(t), p^2(t), \eta^2(t), \mu). \tag{48}$$

The same relation applied to $(\bar{u}, \bar{y})$, the uniform strict complementarity assumption (21) and (A2)–(A3) imply that $\Lambda(0, \bar{u}(t), \bar{y}(t), \bar{p}^2(t), \bar{\eta}^2(t), \bar{\mu}) \leq -C$ for some positive constant C, for all $t \in [\bar{\tau}_{en}, \bar{\tau}_{ex}]$. Therefore, by continuity $\Lambda(0, \bar{u}(t), \bar{y}(t), \bar{p}^2(t), \bar{\eta}^2(t), \bar{\mu}) \leq -C/2$ for all $t \in (\bar{\tau}_{en} - \delta, \bar{\tau}_{ex} + \delta) \supset (\tau_{en}, \tau_{ex})$ for $\delta > 0$, $\|u - \bar{u}\|_\infty$ and $\|\mu - \bar{\mu}\|$ small enough. By Prop. 2.6(iii), for small enough $\|u - \bar{u}\|_\infty$ and $\|\mu - \bar{\mu}\|$, (u, y, p^2, η^2) is arbitrarily close to $(\bar{u}, \bar{y}, \bar{p}^2, \bar{\eta}^2)$ in L^∞ and hence $\Lambda(0, u(t), y(t), p^2(t), \eta^2(t), \mu) \leq -C/4$ on (τ_{en}, τ_{ex}). It follows then from (33)–(34) and (48) that $\ddot{\eta}^2$ is uniformly positive over (τ_{en}, τ_{ex}). This achieves the proof of the theorem. □

Remark 3.2. The regularity of the class of perturbations considered (here, satisfying Def. 2.4) is crucial to show the stability of boundary arcs, as it is the case for first-order state constraints (see [7, Th. 2.1]). If the perturbation is not sufficiently smooth, then boundary arcs are not stable, even if the uniform strict complementarity assumption (21) holds, as it is shown in [18, section 2] for a first-order state constraint and a perturbation that goes to zero in the L^2 norm but not in the $W^{1,\infty}$ norm.

4 Instability of nonreducible touch points

Definition 4.1. Let $\bar{\tau}_{to} \in (0, T)$ be a touch point of a stationary point $(\bar{u}, \bar{y})$ of $(\mathcal{P})$, with alternative multipliers $(\bar{p}^2, \bar{\eta}^2)$.

(a) We say that $\bar{\tau}_{to}$ is *reducible*, if (i) $t \mapsto g^{(2)}(\bar{u}(t), \bar{y}(t))$ is continuous at point $\bar{\tau}_{to}$ (which always holds under assumption (A3)) and (ii)

$$g^{(2)}(\bar{u}(\bar{\tau}_{to}), \bar{y}(\bar{\tau}_{to})) < 0. \tag{49}$$

(b) We say that $\bar{\tau}_{to}$ is *essential*, if

$$[\dot{\bar{\eta}}^2(\bar{\tau}_{to})] > 0. \tag{50}$$

Remark 4.2. Using the classical multipliers $(\bar{p}, \bar{\eta})$ associated with $(\bar{u}, \bar{y})$ in (7)–(10) (recall that $\bar{\eta} = \dot{\bar{\eta}}^2$), (50) is equivalent to

$$[\bar{\eta}(\bar{\tau}_{to})] > 0, \tag{51}$$

which is in accordance with the classical definition of essential touch points.

Let $(\bar{u}, \bar{y})$ be a stationary point of $(\mathcal{P})$ satisfying (A2)–(A3). Assume that $(\bar{u}, \bar{y})$ has a reducible touch point $\bar{\tau}_{to}$. Then given a stationary point (u, y) of $(\mathcal{P}^\mu)$ such that $\|\mu - \bar{\mu}\|$ and $\|u - \bar{u}\|_\infty$ are small enough, it is easy to see (see e.g. [5, section 5.1]) that the mapping $t \mapsto g^\mu(y(t))$ attains its maximum over a neighborhood $(\bar{\tau}_{to} - \delta, \bar{\tau}_{to} + \delta)$ of $\bar{\tau}_{to}$, $\delta > 0$, at a unique point τ_{to}. Therefore, if $g^\mu(y(\tau_{to})) = 0$, (u, y) has a unique touch point in $(\bar{\tau}_{to} - \delta, \bar{\tau}_{to} + \delta)$, and if $g^\mu(y(\tau_{to})) < 0$, the state constraint is locally not active in a neighborhood of $\bar{\tau}_{to}$. Moreover, by Prop. 2.6(ii) and relation (11), $\mathrm{d}\dot{\eta}^2$ weakly-* converges in $\mathcal{M}[0, T]$ to $\mathrm{d}\dot{\bar{\eta}}^2$ when $\|\mu - \bar{\mu}\|, \|u - \bar{u}\|_\infty \to 0$. Therefore, if strict complementarity holds at $\bar{\tau}_{to}$, i.e. if $\bar{\tau}_{to}$ is an essential touch point, this implies that for $\delta > 0$, $\|\mu - \bar{\mu}\|$ and $\|u - \bar{u}\|_\infty$ small enough, $(\bar{\tau}_{to} - \delta, \bar{\tau}_{to} + \delta) \cap \operatorname{supp}(\mathrm{d}\dot{\eta}^2) \neq \emptyset$. Hence by (18) we necessarily have $g^\mu(y(\tau_{to})) = 0$, i.e. τ_{to} is a (essential) touch point of (u, y).

The above discussion shows that touch points that are both reducible and essential are stable. When strict complementarity does not hold, there are two possibilities for nonessential reducible touch points: either the state constraint of the perturbed problem is not active on a neighborhood of $\bar{\tau}_{to}$, or it is active in a neighborhood of $\bar{\tau}_{to}$ at a unique touch point, the latter being essential or not.

We see that the reducibility hypothesis (49) excludes other structural changes. In what follows, we release this reducibility hypothesis and and show that two possible changes in the structure of perturbed stationary points may happen in the neighborhood of a nonreducible touch point: The apparition of a boundary arc or the apparition of a second touch point.

Let now $\bar{\tau}_{to}$ be a nonreducible touch point of $(\bar{u}, \bar{y})$, i.e. such that

$$g^{(2)}(\bar{u}(\bar{\tau}_{to}), \bar{y}(\bar{\tau}_{to})) \ = \ 0. \tag{52}$$

We consider the following assumption (compare to (23))

$$\begin{aligned} \frac{\mathrm{d}^3}{\mathrm{d}t^3} g(\bar{y}(t))|_{t=\bar{\tau}_{to}^-} = g^{(3)}(\dot{\bar{u}}(\bar{\tau}_{to}^-), \bar{u}(\bar{\tau}_{to}), \bar{y}(\bar{\tau}_{to})) &\quad > \quad 0, \\ \frac{\mathrm{d}^3}{\mathrm{d}t^3} g(\bar{y}(t))|_{t=\bar{\tau}_{to}^+} = g^{(3)}(\dot{\bar{u}}(\bar{\tau}_{to}^+), \bar{u}(\bar{\tau}_{to}), \bar{y}(\bar{\tau}_{to})) &\quad < \quad 0. \end{aligned} \tag{53}$$

By (35) and (37), the jumps of $g^{(3)}$ and $\dot{\eta}^2$ at a touch point τ_{to} are related by

$$[g^{(3)}(\dot{u}, u, y)(\tau_{to})] \ = \ g_u^{(2)}(u, y)[\dot{u}(\tau_{to})] \ = \ -\frac{(g_u^{(2)}(u, y))^2}{\tilde{H}_{uu}(u, y, p^2, \eta^2)}[\dot{\eta}^2(\tau_{to})] \ \leq \ 0, \tag{54}$$

where we have $[\dot{\eta}^2(\tau_{to})] = [\eta(\tau_{to})]$ by (11). Therefore, if (53) holds, this implies by (A2)–(A3) that $[\dot{\bar{\eta}}^2(\bar{\tau}_{to})] > 0$. We obtain then the following result.

Lemma 4.3. *Let $(\bar{u}, \bar{y})$ be a stationary point of $(\mathcal{P})$ satisfying (A2)–(A3) and having a nonreducible touch point $\bar{\tau}_{to} \in (0, T)$ satisfying* (53). *Then $\bar{\tau}_{to}$ is an essential touch point, i.e. satisfies* (50).

Let (u, y) be a stationary point of $(\mathcal{P}^\mu)$, with $\|\mu - \bar{\mu}\|$ and $\|u - \bar{u}\|_\infty$ arbitrarily small. We use the notations (41)–(43). At a nonreducible touch point $\bar{\tau}_{to}$ of $(\bar{u}, \bar{y})$, we cannot ensure that the state constraint of the perturbed problem $g(t) := g^\mu(y(t))$ will have a unique maximum point in a neighborhood $(\bar{\tau}_{to} - \delta, \bar{\tau}_{to} + \delta)$ of $\bar{\tau}_{to}$, for small $\delta > 0$.

So let us assume that $g(t)$ has either a boundary arc or an interior arc included in $(\bar{\tau}_{to} - \delta, \bar{\tau}_{to} + \delta)$. We deduce in both cases the existence of a time $t_m \in (\bar{\tau}_{to} - \delta, \bar{\tau}_{to} + \delta)$ where $g^{(2)}$ is maximum (similar to the proof of Th. 3.1) such that

$$g^{(3)}(t_m) = 0.$$

For all $\delta > 0$, if $\|\mu - \bar{\mu}\|$ and $\|u - \bar{u}\|_\infty$ are small enough, then $I(g^\mu(y)) \subset I_\delta(g(\bar{y}))$. Letting $\delta \downarrow 0$, we obtain that $t_m \to \bar{\tau}_{to}$ when $\|\mu - \bar{\mu}\| \to 0$ and $\|u - \bar{u}\|_\infty \to 0$. Hence, (40) implies that when $\|\mu - \bar{\mu}\| \to 0$ and $\|u - \bar{u}\|_\infty \to 0$, using Prop. 2.6(iii),

$$\begin{aligned} g^{(4)}(t_m) + \frac{(g_u^{(2)}(t_m))^2}{\tilde{H}_{uu}(t_m)} \ddot{\eta}^2(t_m) \quad &= \quad -\Lambda(0, u(t_m), y(t_m), p^2(t_m), \eta^2(t_m), \mu) \\ &\to \quad -\Lambda(0, \bar{u}(\bar{\tau}_{to}), \bar{y}(\bar{\tau}_{to}), \bar{p}^2(\bar{\tau}_{to}), \bar{\eta}^2(\bar{\tau}_{to}), \bar{\mu}). \end{aligned} \tag{55}$$

Therefore, if (u,y) has a boundary arc in $(\bar\tau_{to}-\delta,\bar\tau_{to}+\delta)$, we have that $g^{(4)}(t_m)=0$ and $\ddot\eta^2(t_m)\geq 0$, which implies that

$$\Lambda(0,\bar u(\bar\tau_{to}),\bar y(\bar\tau_{to}),\bar p^2(\bar\tau_{to}),\bar\eta^2(\bar\tau_{to}),\bar\mu)\ \leq\ 0. \tag{56}$$

If (u,y) has an interior arc in $(\bar\tau_{to}-\delta,\bar\tau_{to}+\delta)$, then $g^{(4)}(t_m)\leq 0$ (this was shown in the proof of Th. 3.1, recall (47)) and $\ddot\eta^2(t_m)=0$. This implies that

$$\Lambda(0,\bar u(\bar\tau_{to}),\bar y(\bar\tau_{to}),\bar p^2(\bar\tau_{to}),\bar\eta^2(\bar\tau_{to}),\bar\mu)\ \geq\ 0. \tag{57}$$

Conversely, if (56) holds with a strict inequality, then for $\|\mu-\bar\mu\|$ and $\|u-\bar u\|_\infty$ small enough, $g^{(4)}(t_m)+\frac{(g_u^{(2)}(t_m))^2}{\tilde H_{uu}(t_m)}\ddot\eta^2(t_m)>0$, excluding the possibility of an interior arc. Similarly, if (57) holds with a strict inequality, this excludes the possibility of a boundary arc. Using the above arguments, we are able to obtain the following result.

Theorem 4.4. *Let $(\bar u,\bar y)$ be a stationary point of $(\mathcal{P})$ satisfying (A2)–(A3). Assume that $(\bar u,\bar y)$ has a nonreducible and essential touch point $\bar\tau_{to}\in(0,T)$. Set*

$$\bar\lambda(\bar\tau_{to})\ :=\ \Lambda(0,\bar u(\bar\tau_{to}),\bar y(\bar\tau_{to}),\bar p^2(\bar\tau_{to}),\bar\eta^2(\bar\tau_{to}),\bar\mu). \tag{58}$$

Then, for every stable extension $(\mathcal{P}^\mu)$ and for all $\delta>0$ small enough, there exist $\rho,\varrho>0$ such that:
(i) *If $\bar\lambda(\bar\tau_{to})<0$ holds, then all stationary points (u,y) of the perturbed problem $(\mathcal{P}^\mu)$ with $\|\mu-\bar\mu\|<\varrho$ and $\|u-\bar u\|_\infty<\rho$ have either a single touch point τ_{to} or a single boundary arc (τ_{en},τ_{ex}) in $(\bar\tau_{to}-\delta,\bar\tau_{to}+\delta)$. Moreover, in case of a boundary arc (τ_{en},τ_{ex}), (u,y) satisfies the uniform strict complementarity assumption* (21) *on (τ_{en},τ_{ex}).*
(ii) *If $\bar\lambda(\bar\tau_{to})>0$ holds, then all stationary points (u,y) of the perturbed problem $(\mathcal{P}^\mu)$ with $\|\mu-\bar\mu\|<\varrho$ and $\|u-\bar u\|_\infty<\rho$ have either one or two touch points in $(\bar\tau_{to}-\delta,\bar\tau_{to}+\delta)$ and no boundary arc.*

Remark 4.5. Under the assumptions of the above theorem, if $\bar\lambda(\bar\tau_{to})=0$ holds, then we cannot conclude and any structure in the neighborhood of $\bar\tau_{to}$ is a priori possible for a stationary point (u,y) of the perturbed problem $(\mathcal{P}^\mu)$, however small $\|u-\bar u\|_\infty$ and $\|\mu-\bar\mu\|$ are (see Example 4.6 below).

Proof of Th. 4.4. Note first that since $\bar\tau_{to}$ is essential, it follows from Prop. 2.6(ii) and Lemma 2.7 that for $\delta>0$ and $\|u-\bar u\|_\infty$ and $\|\mu-\bar\mu\|$ small enough, $I(g^\mu(y))\cap(\bar\tau_{to}-\delta,\bar\tau_{to}+\delta)$ is not empty. In view of what precedes, it remains to show in the case (ii) when $\bar\lambda(\bar\tau_{to})>0$ that (u,y) cannot have more than one interior arc included in $(\bar\tau_{to}-\delta,\bar\tau_{to}+\delta)$. Since boundary arcs are not possible either, this will show that the only two possibilities for (u,y) is to have one or two touch points in $(\bar\tau_{to}-\delta,\bar\tau_{to}+\delta)$.

If $\bar\lambda(\bar\tau_{to})>0$, then we see by (55) that on an interior arc included in $(\bar\tau_{to}-\delta,\bar\tau_{to}+\delta)$, for $\|u-\bar u\|_\infty$ and $\|\mu-\bar\mu\|$ small enough, for t in the interior arc, the functions (of time) $(g^\mu)^{(4)}(\ddot u,\dot u,u,y)$ being Lipschitz continuous on interior arcs by Lemma 2.1, uniformly w.r.t. μ by Definition 2.4 of a stable extension,

$$g^{(4)}(t)\leq -\tfrac{1}{2}\bar\lambda(\bar\tau_{to})<0,$$

and hence $g^{(3)}$ is strictly decreasing along an interior arc. In addition, $g^{(3)}$ vanishes at some point t_m on the interior of an interior arc where $g^{(2)}$ is maximum and satisfying (47). Now assume that (u,y) has two interior arcs in $(\bar\tau_{to}-\delta,\bar\tau_{to}+\delta)$, say (τ_1,τ_2) and (τ_2,τ_3). Since $g^{(3)}$ is strictly decreasing on the interior arcs and vanishes at an interior point of these arcs, this implies that $g^{(3)}(\tau_2^-)<0$ and $g^{(3)}(\tau_2^+)>0$, and hence, $[g^{(3)}(\tau_2)]>0$. But at the touch point τ_2, $[g^{(3)}(\tau_2)]\leq 0$ by (54), which gives the desired contradiction and shows that (u,y) can only have a single interior arc in $(\bar\tau_{to}-\delta,\bar\tau_{to}+\delta)$, for small enough $\|u-\bar u\|_\infty$ and $\|\mu-\bar\mu\|$ and $\delta>0$.

We end the proof by checking that in the case (i), uniform strict complementarity holds on the boundary arc (τ_{en},τ_{ex}). By (40) and (33), for all t in boundary arc (τ_{en},τ_{ex}) we have that

$$\ddot\eta^2(t)=-\frac{\tilde H_{uu}(t)}{(g_u^{(2)}(t))^2}\Lambda(0,u(t),y(t),p^2(t),\eta^2(t),\mu). \tag{59}$$

Since $c := \bar{\lambda}(\bar{\tau}_{to}) < 0$, it follows that for $\delta > 0$ small enough, $\Lambda(0, \bar{u}(t), \bar{y}(t), \bar{p}^2(t), \bar{\eta}^2(t)) < c/2 < 0$ on $(\bar{\tau}_{to} - \delta, \bar{\tau}_{to} + \delta)$. For $\|u - \bar{u}\|_\infty$ and $\|\mu - \bar{\mu}\|$ small enough, (u, y, p^2, η^2) is arbitrarily close to $(\bar{u}, \bar{y}, \bar{p}^2, \bar{\eta}^2)$ in L^∞ by Prop. 2.6(iii), so if (u, y) has a boundary arc $(\tau_{en}, \tau_{ex}) \subset (\bar{\tau}_{to} - \delta, \bar{\tau}_{to} + \delta)$, we deduce that $\Lambda(0, u(t), y(t), p^2(t), \eta^2(t), \mu) \leq c/4 < 0$ on (τ_{en}, τ_{ex}). With (34)–(33) and (59) this shows that $\ddot{\eta}^2$ is uniformly positive on (τ_{en}, τ_{ex}). This achieves the proof of the theorem. □

Example 4.6. Consider the problem below:

$$\min_{(u,y)\in\mathcal{U}\times\mathcal{Y}} \int_0^1 \left(\frac{u(t)^2}{2} + \mu_1 y_1(t) \right) \mathrm{d}t$$

subject to the dynamics and boundary conditions[1]

$$\begin{aligned} \dot{y}_1(t) &= y_2(t), \quad \dot{y}_2(t) = u(t), && (60)\\ y_1(0) &= y_1(1) = 0, \quad \dot{y}_1(0) = 1 = -\dot{y}_2(1) && (61) \end{aligned}$$

and second-order state constraint

$$y_1(t) \leq \mu_2.$$

The perturbation parameter is $(\mu_1, \mu_2) \in \mathbb{R} \times \mathbb{R}_+^*$. The above problem was studied in [10] for $\mu_1 = 0$ and in [3] for $\mu_1 \neq 0$. By convexity, the first-order optimality condition is necessary and sufficient and the problem has a unique optimal solution.

For the unconstrained problem, the optimality condition reduces to $y_1^{(4)} \equiv -\mu_1$, together with the boundary conditions (61). Therefore the unconstrained optimal trajectory is given by

$$y_1^{uncons}(t) = -\frac{\mu_1}{24}t^4 + \frac{\mu_1}{12}t^3 - \left(1 + \frac{\mu_1}{24}\right)t^2 + t.$$

Its derivatives being given by $\dot{y}_1^{uncons}(t) = \left(t - \frac{1}{2}\right)\left(-\frac{\mu_1}{6}t^2 + \frac{\mu_1}{6}t - 2\right)$ and $\ddot{y}_1^{uncons}(t) = \frac{\mu_1}{2}t(1-t) - 2 - \frac{\mu_1}{12}$, this fourth-order polynomial has on $[0,1]$ a maximum at $t = \frac{1}{2}$ for $\mu_1 \leq 48$, and one local minimum at $t = \frac{1}{2}$ and two maxima, one in $(0, \frac{1}{2})$ and the other in $(\frac{1}{2}, 1)$, for $\mu > 48$. For $\mu_1 \leq 48$ and $\mu_2 = y_1^{uncons}(\frac{1}{2}) = \frac{1}{4} - \frac{\mu_1}{384}$, we have therefore a nonessential touch point at $\tau_{to} = \frac{1}{2}$, which is reducible for $\mu_1 < 48$.

In the sequel we shall consider the case when $\mu_1 < 48$. When μ_2 decreases beyond the value $\frac{1}{4} - \frac{\mu_1}{384}$, the optimal trajectory has one touch point at $\tau_{to} = \frac{1}{2}$ and is given by

$$y_1^{onetouch}(t) = \begin{cases} -\frac{\mu_1}{24}t^4 + \frac{a}{6}t^3 + \frac{b}{2}t^2 + t & \text{on } [0, \frac{1}{2}] \\ -\frac{\mu_1}{24}(t-1)^4 - \frac{a}{6}(t-1)^3 + \frac{b}{2}(t-1)^2 - (t-1) & \text{on } [\frac{1}{2}, 1] \end{cases}$$

with $a = 24 + \frac{\mu_1}{4} - 96\mu_2$ and $b = -8 - \frac{\mu_1}{48} + 24\mu_2$. This touch point becomes nonreducible when $\ddot{y}_1^{onetouch}(\tau_{to}) = 0$ i.e. when $\mu_2 = \frac{1}{6} - \frac{\mu_1}{1152}$, and satisfies (53).

So let us compute the term (58) at the optimal trajectory for a given value of $\bar{\mu}_1 \in (-\infty, 48)$ and $\bar{\mu}_2 := \frac{1}{6} - \frac{\bar{\mu}_1}{1152}$. We have that

$$g(y) = y_1 - \mu_2, \quad g^{(1)}(y) = y_2, \quad g^{(2)}(u, y) = u, \quad g^{(3)}(\dot{u}, u, y) = \dot{u}, \quad g^{(4)}(\ddot{u}, \dot{u}, u, y) = \ddot{u}.$$

The alternative Hamiltonian (14) is given by

$$\tilde{H}^\mu(u, y, p^2, \eta^2) = \frac{u^2}{2} + \mu_1 y_1 + p_1^2 y_2 + p_2^2 u + \eta^2 u$$

and the costate and control equations (16) and (17) are given by

$$\begin{aligned} -\dot{p}_1^2 &= \mu_1, \qquad -\dot{p}_2^2 = p_1^2, \\ 0 &= u + p_2^2 + \eta^2. \end{aligned}$$

[1] Extension of the results of this paper when there are constraints on the final and/or the initial state is possible if a strong controllability condition is assumed, see [6, Section 8].

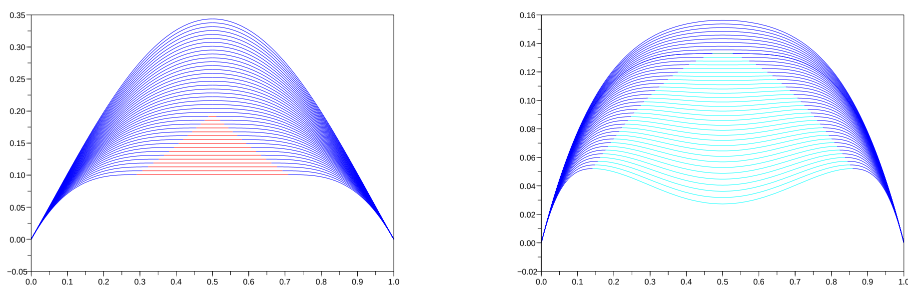
Differentiating twice the last above relation, we obtain

$$0 = \ddot{u} + \mu_1 + \ddot{\eta}^2 = g^{(4)} + \mu_1 + \ddot{\eta}^2.$$

Identifying with (40), we simply have that $\Lambda(g^{(3)}, u, y, p^2, \eta^2, \mu) = \mu_1$, and hence, at the nonreducible touch point $\bar{\tau}_{to} = \frac{1}{2}$,

$$\bar{\lambda}(\bar{\tau}_{to}) = \bar{\mu}_1.$$

Conditions (i) and (ii) of Th. 4.4 are satisfied respectively for $\bar{\mu}_1 < 0$ and for $\bar{\mu}_1 > 0$ (see figure 1 below). Therefore, for $\mu_2 < \frac{1}{6} - \frac{\mu_1}{1152}$, the touch point turns into two touch points if $\mu_1 > 0$ and turns into a boundary



(a) State constraint for $\bar{\mu}_1 = -36$ and varying μ_2. (b) State constraint for $\bar{\mu}_1 = 36$ and varying μ_2.

Figure 1: Transformation of a nonreducible touch point into a boundary arc or into two touch points for $\bar{\mu}_1 \neq 0$ when μ_2 decreases.

arc if $\mu_1 < 0$, and strict complementarity holds on that boundary arc since $\ddot{\eta}^2 \equiv -\mu_1 > 0$.

If $\bar{\mu}_1 = 0$, then $\bar{\lambda}(\bar{\tau}_{to}) = 0$ and we cannot conclude for the structure of the solutions of the perturbed problem. For $\mu_2 < \frac{1}{6}$ and $\mu_1 = 0$, a boundary arc appears but strict complementarity does not hold on that boundary arc since $\ddot{\eta}^2 \equiv -\mu_1 = 0$. If we take e.g. $\mu_2 = \frac{1}{6} - \frac{\mu_1}{1152} - \varepsilon\mu_1^2$, with $\varepsilon > 0$ a fixed parameter, we have in the neighborhood of the nonreducible touch point $\bar{\tau}_{to}$ a boundary arc for $\mu_1 < 0$ and two touch points for $\mu_1 > 0$.

5 Stability analysis

Let $(\bar{u}, \bar{y})$ be a stationary point of $(\mathcal{P})$ with alternative multipliers $(\bar{p}^2, \bar{\eta}^2)$. Let $\mathcal{V} := L^2(0,T)$. For $v \in \mathcal{V}$, recall that we denote by z_v the unique solution in $H^1(0,T;\mathbb{R}^n)$ of the linearized state equation (30). The quadratic form involved in the second-order optimality conditions in [16] is as follows: For $v \in \mathcal{V}$,

$$\mathcal{Q}(v) := \int_0^T D^2_{(u,y)(u,y)}\tilde{H}(\bar{u}, \bar{y}, \bar{p}^2, \bar{\eta}^2)((v, z_v), (v, z_v))\mathrm{d}t \; + \; \phi_{yy}(\bar{y}(T))(z_v(T), z_v(T)). \tag{62}$$

The *extended critical cone* used in the stability analysis is defined as the set of $v \in \mathcal{V}$ such that

$$g_y(\bar{y}(t))z_v(t) = 0 \qquad \text{for all } t \in \operatorname{supp}(\mathrm{d}\dot{\bar{\eta}}^2). \tag{63}$$

This set is obtained from the classical critical cone, defined as the set of $v \in \mathcal{V}$ satisfying (63) and

$$g_y(\bar{y}(t))z_v(t) \leq 0 \qquad \text{for all } t \in I(g(\bar{y})) \setminus \operatorname{supp}(\mathrm{d}\dot{\bar{\eta}}^2), \tag{64}$$

by omission of the inequality constraint (64). The strong second-order sufficient condition used in the stability analysis is:

$$\mathcal{Q}(v) > 0, \quad \text{for all } v \in \mathcal{V},\ v \neq 0, \text{ satisfying (63)}. \tag{65}$$

This condition is a natural strengthening of the second-order sufficient condition of [5, Th. 18]

$$\mathcal{Q}(v) > 0, \quad \text{for all } v \in \mathcal{V},\ v \neq 0, \text{ satisfying (63)-(64)}. \tag{66}$$

The strengthened Legendre-Clebsch condition (20) implies that the quadratic form $\mathcal{Q}$ is a *Legendre form*, i.e. a weakly lower semi-continuous quadratic form with the property that if a sequence v_n weakly converges to v in L^2 and if $\mathcal{Q}(v_n) \to \mathcal{Q}(v)$, then v_n converges to v strongly in L^2. Consequently, (65) (resp. (66)) is equivalent to the existence of some $c > 0$ such that $\mathcal{Q}(v) \geq c\|v\|_2^2$ for all $v \in \mathcal{V}$ satisfying (63) (resp. satisfying (63)-(64)).

For first-order state constraints, the stability analysis for the homotopy algorithm in [7] was conducted using a shooting approach. For second-order state constraints, a shooting approach can be used for the stability analysis if all the touch points are reducible, see [20, 8], but not in presence of nonreducible touch points, since in that case the structure is not stable by Th. 4.4. For this reason, a stability result has been obtained in [16] (Th. 5.1 below) that makes no assumptions on the structure of the trajectory, and hence applies when the structure of the trajectory is not stable. This result is based on a variant of Robinson's strong regularity theory [25] and extends the stability results known for first-order state constraints, see [12, 17].

Theorem 5.1 ([16, Th. 4.3])**.** *Let $(\bar{u}, \bar{y})$ be a local solution of $(\mathcal{P})$, satisfying (A2)–(A3) and the strong second-order sufficient condition* (65)*, and let $(\mathcal{P}^\mu)$ be a stable extension of $(\mathcal{P})$. Then there exist $c, \rho, \kappa, \tilde{\kappa} > 0$ and a neighborhood $\mathcal{N}$ of $\bar{\mu}$, such that for all $\mu \in \mathcal{N}$, $(\mathcal{P}^\mu)$ has a unique stationary point (u^μ, y^μ) with $\|u^\mu - \bar{u}\|_\infty < \rho$ and unique associated alternative multipliers $(p^{2,\mu}, \eta^{2,\mu})$, and for all $\mu, \mu' \in \mathcal{N}$,*

$$\begin{aligned} \|u^\mu - u^{\mu'}\|_2,\ \|y^\mu - y^{\mu'}\|_{1,2},\ \|p^{2,\mu} - p^{2,\mu'}\|_{1,2},\ \|\eta^{2,\mu} - \eta^{2,\mu'}\|_2 &\leq \kappa\|\mu - \mu'\|, && (67)\\ \|u^\mu - u^{\mu'}\|_\infty, \|y^\mu - y^{\mu'}\|_{1,\infty}, \|p^{2,\mu} - p^{2,\mu'}\|_{1,\infty}, \|\eta^{2,\mu} - \eta^{2,\mu'}\|_\infty &\leq \tilde{\kappa}\|\mu - \mu'\|^{2/3}. && (68)\end{aligned}$$

Moreover, (u^μ, y^μ) is a local solution of $(\mathcal{P}^\mu)$ satisfying the uniform quadratic growth condition (28) *and the strong second-order sufficient condition* (65)*.*

Proof. The theorem follows from [16, Th. 4.3], excepted for the fact that (u^μ, y^μ) satisfies the strong second-order sufficient condition (65). The latter can be proved by contradiction, by a slight modification of the proof of [16, Prop. 4.2], using Prop. 2.6, Lemma 2.7, and the fact that $\mathcal{Q}$ is a Legendre form. □

6 The shooting algorithm

By Th. 5.1, the perturbed problem $(\mathcal{P}^\mu)$ has a locally unique local solution. The objective of this section is to see, under additional assumptions, how we could use the shooting algorithm and the results of Theorems 3.1 and 4.4 to obtain in practice in the homotopy algorithm the solution of the perturbed problem.

Let us first recall the shooting algorithm for a second-order scalar state constraint (see [10, 23, 20, 8]). The alternative multipliers used in the shooting algorithm are denoted by (p_2, η_2), with the '2' as subscript, not to be confused with the multipliers (p^2, η^2) (with the '2' as superscript) used in the stability analysis. Let us recall that the multipliers used in the shooting algorithm (p_2, η_2) are defined, on each boundary arc (τ_{en}, τ_{ex}) of the trajectory, by

$$\begin{aligned} \eta_1(t) &:= \int_{(t,\tau_{ex}]} \mathrm{d}\eta(s) = \eta(\tau_{ex}^+) - \eta(t^+), \qquad \eta_2(t) := \int_t^{\tau_{ex}} \eta_1(s)\mathrm{d}s, && (69)\\ p_2(t) &:= p(t) - \eta_1(t)g_y(y(t)) - \eta_2 g_y^{(1)}(y(t)) && (70)\end{aligned}$$

and $\eta_1(t), \eta_2(t), p_2(t) = 0$ outside boundary arcs. Here p and η denote the multipliers associated with a stationary point (u, y) in the classical optimality condition (7)–(10).

Why do we use so many different multipliers? The multipliers η^2, p^2 are very useful in the stability analysis because they are continuous and converge uniformly. The multipliers (p_2, η_2) used in the shooting algorithm have jumps, and these jumps are used as additional degrees of freedom in the shooting algorithm, in order to have as many free parameters as conditions to satisfy. An explicit relation between these multipliers (p_2, η_2) and (p^2, η^2) is made precise later, see (113)–(115).

Let $(\bar{u}, \bar{y})$ be a stationary point of $(\mathcal{P})$ satisfying (A2)–(A3) and the assumption below:

(A4) $(\bar u, \bar y)$ has finitely many boundary arcs and finitely many touch points and the state constraint is not active at final time, i.e. $g(\bar y(T)) < 0$.

Denote by $\bar{\mathcal{T}}_{en}$, $\bar{\mathcal{T}}_{ex}$ and $\bar{\mathcal{T}}_{to}$ the (finite and possibly empty) sets of respectively entry, exit and touch times of the trajectory $(\bar u, \bar y)$, and its set of junction points by $\bar{\mathcal{T}} := \bar{\mathcal{T}}_{en} \cup \bar{\mathcal{T}}_{ex} \cup \bar{\mathcal{T}}_{to}$. Let $N_{ba} := |\bar{\mathcal{T}}_{en}| = |\bar{\mathcal{T}}_{ex}|$ and $N_{to} := |\bar{\mathcal{T}}_{to}|$. Moreover let us introduce the following notation. Given a real-valued function φ over $[0,T]$ and a finite subset $\mathcal{S}$ of $(0,T)$, assuming w.l.o.g. the elements of $\mathcal{S}$ in increasing order, we may define $\varphi(\mathcal{S}) := (\varphi(\tau))_{\tau\in\mathcal{S}} \in \mathbb{R}^{\mathrm{Card}\,\mathcal{S}}$. We adopt a similar convention for vectors and define $\nu_{\mathcal{S}} := (\nu_\tau)_{\tau\in\mathcal{S}} \in \mathbb{R}^{\mathrm{Card}\,\mathcal{S}}$.

The shooting algorithm is as follows. The unknown are the initial value of the costate p_0, the (finite) sets of entry, exit and touch points of the trajectory, respectively $\mathcal{T}_{en}$, $\mathcal{T}_{ex}$ and $\mathcal{T}_{to}$, and the jump parameters of the costate. More precisely, there are two jump parameters $\nu^1_{\tau_{en}}$ and $\nu^2_{\tau_{en}}$ for each entry point $\tau_{en} \in \mathcal{T}_{en}$ and one jump parameter $\nu_{\tau_{to}}$ for each touch point $\tau_{to} \in \mathcal{T}_{to}$. The shooting mapping $\mathcal{F}$ in a neighborhood of $(\bar u, \bar y)$ is defined by

$$\begin{array}{rcl} \mathcal{F} : \mathbb{R}^n \times (\mathbb{R}^{N_{ba}})^4 \times (\mathbb{R}^{N_{to}})^2 & \rightarrow & \mathbb{R}^n \times (\mathbb{R}^{N_{ba}})^4 \times (\mathbb{R}^{N_{to}})^2, \\ \begin{pmatrix} p_0 \\ \nu^1_{\mathcal{T}_{en}} \\ \nu^2_{\mathcal{T}_{en}} \\ \mathcal{T}_{en} \\ \mathcal{T}_{ex} \\ \nu_{\mathcal{T}_{to}} \\ \mathcal{T}_{to} \end{pmatrix} & \mapsto & \begin{pmatrix} p_2(T) - \phi_y(y(T)) \\ g(y(\mathcal{T}_{en})) \\ g^{(1)}(y(\mathcal{T}_{en})) \\ g^{(2)}(u(\mathcal{T}_{en}^-), y(\mathcal{T}_{en})) \\ g^{(2)}(u(\mathcal{T}_{ex}^+), y(\mathcal{T}_{ex})) \\ g(y(\mathcal{T}_{to})) \\ g^{(1)}(y(\mathcal{T}_{to})) \end{pmatrix} \end{array}$$

where (u, y, p_2, η_2) are the solution of:

$$\begin{aligned} \dot y &= f(u,y) \quad \text{on } [0,T], \quad y(0) = y_0 && (71)\\ -\dot p_2 &= \tilde H_y(u,y,p_2,\eta_2) \quad \text{on } [0,T]\setminus\mathcal{T}, \quad p_2(0) = p_0, && (72)\\ 0 &= \tilde H_u(u,y,p_2,\eta_2) \quad \text{on } [0,T]\setminus\mathcal{T}, && (73)\\ 0 &= g^{(2)}(u,y) \quad \text{on boundary arcs} && (74)\\ 0 &= \eta_2 \quad \text{on interior arcs} && (75)\\ [p_2(\tau_{en})] &= -\nu^1_{\tau_{en}} g_y(y(\tau_{en})) - \nu^2_{\tau_{en}} g^{(1)}_y(y(\tau_{en})) \quad \text{at entry times } \tau_{en} \in \mathcal{T}_{en} && (76)\\ [p_2(\tau_{to})] &= -\nu_{\tau_{to}} g_y(y(\tau_{to})) \quad \text{at touch points } \tau_{to} \in \mathcal{T}_{to}. && (77) \end{aligned}$$

A vector of shooting parameters will be denoted by θ. With a stationary point of $(\mathcal{P})$ satisfying (A2)–(A4) is associated a unique set of shooting parameters, which is a zero of the shooting mapping. The vector of shooting parameters of $(\bar u, \bar y)$ will be denoted by $\bar\theta$. More generally the 'bar' will refer in what follows to shooting parameters associated with the reference trajectory $(\bar u, \bar y)$. Let us recall (see [8, Rem. 2.11(ii)]) that using the multipliers $(\bar p_2, \bar\eta_2)$ uniquely associated with $(\bar u, \bar y)$ in the shooting algorithm, assumption (A3) is equivalent to

$$\begin{aligned} &\bar u \text{ is continuous over } [0,T] \text{ and} \\ &\exists\, \alpha > 0, \quad \tilde H_{uu}(\bar u(t), \bar y(t), \bar p_2(t^\pm), \bar\eta_2(t^\pm)) \geq \alpha \quad \text{for all } t \in [0,T]. \end{aligned} \tag{78}$$

If (u,y) is a trajectory associated with a zero of the shooting mapping, with alternative shooting multipliers (p_2, η_2), then u, p_2 and η_2 are piecewise continuous on $[0,T]$ and have their set of discontinuity times included in the set of junction times $\mathcal{T} := \mathcal{T}_{en} \cup \mathcal{T}_{ex} \cup \mathcal{T}_{to}$. Let us recall the additional conditions that are automatically satisfied by a zero of the shooting mapping and the additional conditions, under which a zero of the shooting mapping is associated with a stationary point of the optimal control problem. Given $a, b \in \mathbb{R}$, set $[a,b] := \{(1-\lambda)a + \lambda b\ ;\ \lambda \in [0,1]\}$.

Lemma 6.1 ([8, Prop. 2.15 and Rem. 2.16])**.** *Let (u,y) be the trajectory associated with a zero of the shooting mapping, with alternative shooting multipliers (p_2, η_2). Assume that there exists $\beta, \alpha > 0$ such that*

$$\begin{aligned} \beta &\leq |g^{(2)}_u(\hat u, y(t))| \quad \textit{for all } \hat u \in [u(t^-), u(t^+)] \textit{ and all } t \in I(g(y)); && (79)\\ \alpha &\leq \tilde H_{uu}(\hat u, y(t), p_2(t^\pm), \eta_2(t^\pm)) \quad \textit{for all } \hat u \in [u(t^-), u(t^+)] \textit{ and all } t \in [0,T]. && (80) \end{aligned}$$

Then: (i) u *is continuous over* $[0,T]$.
(ii) *For each boundary arc* (τ_{en}, τ_{ex}) *of* (u,y)*, the following holds:*

$$\eta_2(\tau_{en}^+) = \nu_{\tau_{en}}^2 \qquad \text{and} \qquad \eta_2(\tau_{ex}^-) = 0. \tag{81}$$

Proposition 6.2 ([8, Corollary 2.17])**.** *A zero of the shooting mapping is associated with a stationary point* (u,y) *of* $(\mathcal{P})$ *satisfying (A2),* (78)*, and (A4), with alternative shooting multipliers* (p_2,η_2)*, iff:*

$$\begin{aligned} g(y(t)) &\le 0 \quad \text{on interior arcs,} && (82)\\ 0 &\le \ddot{\eta}_2(t) \quad \text{on boundary arcs,} && (83)\\ 0 &\le \nu_{\tau_{en}}^1 + \dot{\eta}_2(\tau_{en}^+) \quad \text{for each entry point } \tau_{en}, && (84)\\ \dot{\eta}_2(\tau_{ex}^-) &\le 0 \quad \text{for each exit point } \tau_{ex} && (85)\\ 0 &\le \nu_{\tau_{to}} \quad \text{for each touch point } \tau_{to}. && (86) \end{aligned}$$

Lemma 6.3. *Let* (u,y) *be the trajectory associated with a zero of the shooting mapping satisfying (A2),* (78)*, and (A4). Then the additional conditions* (84) *and* (85) *are equivalent, respectively, to*

$$g^{(3)}(\dot{u}(\tau_{en}^-), u(\tau_{en}), y(\tau_{en})) \ge 0 \qquad \text{and} \qquad g^{(3)}(\dot{u}(\tau_{ex}^+), u(\tau_{ex}), y(\tau_{ex})) \le 0 \tag{87}$$

where the function $g^{(3)}$ *is defined by* (35)*.*

Proof. By time differentiation of (73) on the interior of arcs, we have (omitting the arguments (u,y,p_2,η_2))

$$0 \;=\; \tilde{H}_{uu}\dot{u} + \tilde{H}_{uy}f - \tilde{H}_y f_u + \dot{\eta}_2 g_u^{(2)}. \tag{88}$$

Taking the jumps at entry time τ_{en}, we have by (76) and (81) (omitting arguments)

$$\begin{aligned} [\tilde{H}_{uu}] &= [p_2]f_{uu} + [\eta_2]g_{uu}^{(2)} \;=\; -\nu_{\tau_{en}}^1 g_y f_{uu} - \nu_{\tau_{en}}^2 g_y^{(1)} f_{uu} + \nu_{\tau_{en}}^2 g_{uu}^{(2)}\\ &= -\nu_{\tau_{en}}^1 g_{uu}^{(1)} - \nu_{\tau_{en}}^2 g_{uu}^{(2)} + \nu_{\tau_{en}}^2 g_{uu}^{(2)}\\ &= 0,\\ [\tilde{H}_{uy}]f - [\tilde{H}_y]f_u &= [p_2]f_{uy}f + [\eta_2]g_{uy}^{(2)}f - [p_2]f_y f_u - [\eta_2]g_y^{(2)}f_u\\ &= -\nu_{\tau_{en}}^1(g_y f_{uy} f - g_y f_y f_u) - \nu_{\tau_{en}}^2(g_y^{(1)} f_{uy} f - g_{uy}^{(2)} f - g_y^{(1)} f_y f_u + g_y^{(2)} f_u). \end{aligned}$$

Using that $g_{uy}^{(j)} = g_{yy}^{(j-1)} f_u + g_y^{(j-1)} f_{uy}$, $j=1,2$, that $g_u^{(2)} = g_y^{(1)} f_u = g_{yy} f f_u + g_y f_y f_u$ and that $g_{uy}^{(1)} \equiv 0$, we obtain

$$\begin{aligned} [\tilde{H}_{uy}]f - [\tilde{H}_y]f_u &= -\nu_{\tau_{en}}^1(g_{uy}^{(1)} f - g_{yy} f_u f - g_u^{(2)} + g_{yy} f f_u) - \nu_{\tau_{en}}^2(-g_{yy}^{(1)} f_u f + g_{yy}^{(1)} f f_u)\\ &= \nu_{\tau_{en}}^1 g_u^{(2)}. \end{aligned}$$

Therefore, taking the jump of (88) at τ_{en}, we obtain

$$0 \;=\; \tilde{H}_{uu}[\dot{u}(\tau_{en})] + (\nu_{\tau_{en}}^1 + [\dot{\eta}_2(\tau_{en})])g_u^{(2)}.$$

By (35), we have that $[g^{(3)}(\dot{u}(\tau_{en}), u(\tau_{en}), y(\tau_{en}))] = g_u^{(2)}[\dot{u}(\tau_{en})]$ and hence, since $g^{(3)}$ vanishes on the interior of the boundary arc,

$$\nu_{\tau_{en}}^1 + \dot{\eta}_2(\tau_{en}^+) = \frac{\tilde{H}_{uu}}{(g_u^{(2)})^2} g^{(3)}(\dot{u}(\tau_{en}^-), u(\tau_{en}), y(\tau_{en})). \tag{89}$$

Since $\tilde{H}_{uu}/(g_u^{(2)})^2$ is positive by (78) and (A2), the additional condition (84) is equivalent to the first condition of (87). Using similar arguments at exit points, the result follows. ∎

Remark 6.4. It follows from the above lemma that (82) together with the continuity of u imply that (84)–(85) are satisfied, since a Taylor expansion of the state constraint near entry/exit of boundary arcs yields

$$0 \ge g(y(t)) = g^{(3)}(\dot{u}(\tau^\pm), u(\tau), y(\tau))\frac{(t-\tau)^3}{6} + o(|t-\tau|^3),$$

where $\tau^\pm$ stands for τ_{en}^- or τ_{ex}^+, implying (87), and in turn (84)–(85).

In what follows, $(\mathcal{P}^\mu)$ denotes a stable extension of $(\mathcal{P})$, and to indicate the dependence on μ of the data g, f, ℓ, ϕ and $\tilde{H}$, we will denote in what follows the shooting mapping by $\mathcal{F}(\cdot,\mu)$.

6.1 Well-posedness with nonreducible touch points

We assume in addition to (A2)–(A4) that

(A5) The strict complementarity assumption (21) holds on each (regular) boundary arc $(\bar{\tau}_{en}, \bar{\tau}_{ex})$ of $(\bar{u}, \bar{y})$;

(A6) **(i)** Each nonreducible touch point $\bar{\tau}_{to}$ of $(\bar{u}, \bar{y})$ satisfies (53);

(ii) Each nonreducible touch point $\bar{\tau}_{to}$ of $(\bar{u}, \bar{y})$ satisfies $\bar{\lambda}(\bar{\tau}_{to}) < 0$, where $\bar{\lambda}(\bar{\tau}_{to})$ is defined by (58).

Assumption (A6)(i) implies by Lemma 4.3 that all *nonreducible* touch points of $(\bar{u}, \bar{y})$ are *essential*. Therefore, by (A6)(i) all *nonessential* touch points of $(\bar{u}, \bar{y})$ are *reducible*, i.e. satisfy (49).

We exclude in (A6)(ii) the case when $\bar{\lambda}(\bar{\tau}_{to}) = 0$, since in that case, by Remark 4.5, we have no information on the structure of solutions of the perturbed problem, which is not very useful for the homotopy algorithm. We also exclude the case when $\bar{\lambda}(\bar{\tau}_{to}) > 0$, though we know by Th. 4.4 that in that case the solutions of the perturbed problem have either one or two touch points in the neighborhood of $\bar{\tau}_{to}$. The reason to leave aside this case in the following analysis is that singularities happen in the shooting algorithm when a touch point turns into two touch points (this is discussed more precisely in Remark 8.4 at the end of the paper).

Definition 6.5. Let $(\bar{u}, \bar{y})$ be a stationary point of $(\mathcal{P})$ satisfying (A2)–(A6) and let $(\mathcal{P}^\mu)$ be a stable extension of $(\mathcal{P})$. We say that a stationary point (u, y) of $(\mathcal{P}^\mu)$ has a *neighboring structure* to that of $(\bar{u}, \bar{y})$ if there exists a small $\delta > 0$, $\delta < \min_{\tau,\tau' \in \bar{\mathcal{T}}, \tau \neq \tau'} |\tau - \tau'|$, such that (a)–(e) below hold:

(a) The contact set $I(g^\mu(y))$ is included in $I_\delta(g(\bar{y})) = \{t \in [0, T] : \text{dist}\{t, I(g(\bar{y}))\} < \delta\}$;

(b) For each *boundary arc* $(\bar{\tau}_{en}, \bar{\tau}_{ex})$ of $(\bar{u}, \bar{y})$, (u, y) has on $(\bar{\tau}_{en} - \delta, \bar{\tau}_{ex} + \delta)$ a unique boundary arc (τ_{en}, τ_{ex});

(c) For each *essential* and *reducible* touch point $\bar{\tau}_{to}$ of $(\bar{u}, \bar{y})$, (u, y) has on $(\bar{\tau}_{to} - \delta, \bar{\tau}_{to} + \delta)$ a unique touch point τ_{to};

(d) For each *nonessential* touch point $\bar{\tau}_{to}$ of $(\bar{u}, \bar{y})$, either the state constraint $g^\mu(y)$ is not active on $(\bar{\tau}_{to} - \delta, \bar{\tau}_{to} + \delta)$ or (u, y) has on $(\bar{\tau}_{to} - \delta, \bar{\tau}_{to} + \delta)$ a unique touch point τ_{to};

(e) For each *nonreducible* touch point $\bar{\tau}_{to}$ of $(\bar{u}, \bar{y})$, (u, y) has on $(\bar{\tau}_{to} - \delta, \bar{\tau}_{to} + \delta)$ either a unique touch point τ_{to} or a unique boundary arc (τ_{en}, τ_{ex}).

We denote by $\bar{\mathcal{T}}^{ess}_{red}$, $\bar{\mathcal{T}}^{nes}$, and $\bar{\mathcal{T}}_{nrd}$ the sets of respectively essential and reducible, nonessential, and nonreducible touch points of the trajectory $(\bar{u}, \bar{y})$. Set $N_{nes} := |\bar{\mathcal{T}}^{nes}|$ and $N_{nrd} := |\bar{\mathcal{T}}_{nrd}|$. By the above definition, there are $N_s := 2^{N_{nes}+N_{nrd}}$ different neighboring structures to that of $(\bar{u}, \bar{y})$. For $j = 1, \dots, N_s$, denote by $\mathcal{F}_j$ the shooting mappings corresponding to each of those different neighboring structures. For each nonessential touch point $\bar{\tau}_{to}$ of $(\bar{u}, \bar{y})$, the latter is introduced or not in the shooting mapping $\mathcal{F}_j$ (with a zero jump parameter $\bar{\nu}_{\tau_{to}}$), and for each nonreducible touch point $\bar{\tau}_{to}$ of $(\bar{u}, \bar{y})$, the latter is introduced as a touch point or as a boundary arc (of zero length) in the shooting mapping $\mathcal{F}_j$. More precisely, similarly to first-order state constraints (see [7, section 4.2]) since $g^{(2)}(\bar{u}(\bar{\tau}_{to}^\pm), \bar{y}(\bar{\tau}_{to})) = 0$ a nonreducible touch point $\bar{\tau}_{to}$ can be seen as a boundary arc of zero length, by taking

$$\bar{\tau}_{en} := \bar{\tau}_{to} =: \bar{\tau}_{ex} \tag{90}$$

and, in view of the jump conditions (76)–(77),

$$\bar{\nu}^1_{\tau_{en}} := \bar{\nu}_{\tau_{to}} \qquad \text{and} \qquad \bar{\nu}^2_{\tau_{en}} := 0. \tag{91}$$

For $j = 1, \dots, N_s$, denote by $\bar{\theta}_j$ the vector of shooting parameters, of appropriate dimension, associated with $(\bar{u}, \bar{y})$ in the shooting mapping $\mathcal{F}_j$.

For $v \in \mathcal{V}$ in the extended critical cone (i.e. satisfying (63)), we consider the additional constraint below:

$$g^{(1)}_y(y(\bar{\tau}_{to}))z_v(\bar{\tau}_{to}) = 0 \qquad \text{for all } \bar{\tau}_{to} \in \bar{\mathcal{T}}_{nrd}. \tag{92}$$

Recall that z_v is the solution of (30). A sufficient condition ensuring the well-posedness of the shooting algorithm, as we will see, is

$$\mathcal{Q}(v) - \sum_{\tau \in \bar{\mathcal{T}}_{red}^{ess}} \bar{\nu}_\tau \frac{(g_y^{(1)}(\bar{y}(\tau))z_v(\tau))^2}{g^{(2)}(\bar{u}(\tau), \bar{y}(\tau))} > 0, \quad \text{for all } v \in \mathcal{V},\ v \neq 0, \text{ satisfying (63) and (92)}, \tag{93}$$

where $\mathcal{Q}$ is given by (62). Note that the sum in (93) is nonpositive. Therefore, the strong second-order sufficient condition (65) used in the stability analysis implies that the weaker condition (93) is satisfied.

Lemma 6.6. *Let $(\bar{u}, \bar{y})$ be a stationary point of $(\mathcal{P})$ satisfying (A2)–(A6) and (93). Then there exists a neighborhood W of $\bar{\mu}$ and, for each $j = 1, \dots, N_s$, a neighborhood V_j of $\bar{\theta}_j$ such that for each $\mu \in W$, the equation*

$$\mathcal{F}_j(\theta, \mu) = 0 \tag{94}$$

has a unique solution θ_j^μ in V_j, which is C^1 w.r.t. μ.

Of course, if nonreducible touch points are converted into boundary arcs in the shooting mapping $\mathcal{F}_j$, it may happen that for μ in the neighborhood of $\bar{\mu}$, the solution θ_j^μ of (94) is such that some entry times are greater than the corresponding exit times. In that case the trajectory associated with θ_j^μ by (71)–(77) has no physical meaning since not single-valued. In Lemma 6.9 we will give necessary and sufficient conditions so that a solution θ_j^μ of (94) is associated with a stationary point of $(\mathcal{P}^\mu)$.

Proof of Lemma 6.6. We follows the ideas of the proof of [8, Th. 3.3] and include the presence of nonreducible touch points. Let us show that the Jacobian $D_\theta \mathcal{F}_j(\bar{\theta}_j, \bar{\mu})$ is invertible, for all $j = 1, \dots, N_s$. It will then follow from the implicit function theorem that (94) has a locally unique solution for μ in a neighborhood of $\bar{\mu}$ which is C^1 w.r.t. μ.

Let $\mathcal{F}_j$ be one of these shooting mappings. Let $\omega := \left(\pi_0, \gamma^1_{\mathcal{T}_{en}}, \gamma^2_{\mathcal{T}_{ex}}, \sigma_{\mathcal{T}_{en}}, \sigma_{\mathcal{T}_{ex}}, \gamma_{\mathcal{T}_{to}}, \sigma_{\mathcal{T}_{to}}\right)^\top$ be such that $D_\theta \mathcal{F}_j(\bar{\theta}_j, \bar{\mu})\omega = 0$. Then, by differentiation of the shooting mapping, we have

$$\begin{aligned}
0 &= \pi_2(T) - \phi_{yy}(y(T))z(T), && (95)\\
0 &= g_y(\bar{y}(\bar{\tau}_{en}))z(\bar{\tau}_{en}) \quad \text{for all entry points } \bar{\tau}_{en}, && (96)\\
0 &= g_y^{(1)}(\bar{y}(\bar{\tau}_{en}))z(\bar{\tau}_{en}) \quad \text{for all entry points } \bar{\tau}_{en}, && (97)\\
0 &= Dg^{(2)}(\bar{u}(\bar{\tau}_{en}), \bar{y}(\bar{\tau}_{en}))(v(\bar{\tau}_{en}^-), z(\bar{\tau}_{en})) + \sigma_{\bar{\tau}_{en}} \frac{\mathrm{d}}{\mathrm{d}t} g^{(2)}(\bar{u}(t), \bar{y}(t))|_{t=\bar{\tau}_{en}^-} \\
&\quad \text{for all entry points } \bar{\tau}_{en}, && (98)\\
0 &= Dg^{(2)}(\bar{u}(\bar{\tau}_{ex}), \bar{y}(\bar{\tau}_{ex}))(v(\bar{\tau}_{ex}^+), z(\bar{\tau}_{ex})) + \sigma_{\bar{\tau}_{ex}} \frac{\mathrm{d}}{\mathrm{d}t} g^{(2)}(\bar{u}(t), \bar{y}(t))|_{t=\bar{\tau}_{ex}^+} \\
&\quad \text{for all exit points } \bar{\tau}_{ex}, && (99)\\
0 &= g_y(\bar{y}(\bar{\tau}_{to}))z(\bar{\tau}_{to}) \quad \text{for all touch points } \bar{\tau}_{to}, && (100)\\
0 &= g_y^{(1)}(\bar{y}(\bar{\tau}_{to}))z(\bar{\tau}_{to}) + \sigma_{\bar{\tau}_{to}} g^{(2)}(\bar{u}(\bar{\tau}_{to}), \bar{y}(\bar{\tau}_{to})) \quad \text{for all touch points } \bar{\tau}_{to}, && (101)
\end{aligned}$$

where (v, z, π_2, ζ_2) are the solutions of the variational system below (the arguments $(\bar{u}, \bar{y}, \bar{p}_2, \bar{\eta}_2)$ are omitted)

$$\begin{aligned}
\dot{z} &= f_u v + f_y z \quad \text{on } [0,T], \quad z(0) = 0, && (102)\\
-\dot{\pi}_2 &= \tilde{H}_{yu} v + \tilde{H}_{yy} z + \pi_2 f_y + \zeta_2 g_y^{(2)} \quad \text{on } [0,T] \setminus \mathcal{T}, \quad \pi_2(0) = \pi_0, && (103)\\
0 &= \tilde{H}_{uu} v + \tilde{H}_{uy} z + \pi_2 f_u + \zeta_2 g_u^{(2)} \quad \text{on } [0,T] \setminus \mathcal{T}, && (104)\\
0 &= g_u^{(2)} v + g_y^{(2)} z \quad \text{on boundary arcs}, && (105)\\
0 &= \zeta_2 \quad \text{on interior arcs}, && (106)\\
[\pi_2(\bar{\tau}_{en})] &= -\bar{\nu}^1_{\bar{\tau}_{en}} g_{yy}(\bar{y}(\bar{\tau}_{en}))z(\bar{\tau}_{en}) - \bar{\nu}^2_{\bar{\tau}_{en}} g_{yy}^{(1)}(\bar{y}(\bar{\tau}_{en}))z(\bar{\tau}_{en}) - \gamma^1_{\bar{\tau}_{en}} g_y(\bar{y}(\bar{\tau}_{en})) \\
&\quad - (\gamma^2_{\bar{\tau}_{en}} + \sigma_{\bar{\tau}_{en}} \bar{\nu}^1_{\bar{\tau}_{en}}) g_y^{(1)}(\bar{y}(\bar{\tau}_{en})) \quad \text{for all entry points } \bar{\tau}_{en}, && (107)\\
[\pi_2(\bar{\tau}_{to})] &= -\bar{\nu}_{\bar{\tau}_{to}} g_{yy}(\bar{y}(\bar{\tau}_{to}))z(\bar{\tau}_{to}) - \gamma_{\bar{\tau}_{to}} g_y(\bar{y}(\bar{\tau}_{to})) - \sigma_{\bar{\tau}_{to}} \bar{\nu}_{\bar{\tau}_{to}} g_y^{(1)}(\bar{y}(\bar{\tau}_{to})) \\
&\quad \text{for all touch points } \bar{\tau}_{to}. && (108)
\end{aligned}$$

The jump condition of the costate (107) follows from [8, Lemma 3.7]. Recall that for nonreducible touch points $\bar\tau_{to} = \bar\tau_{en}$ converted into a boundary arc in $\mathcal{F}_j$, we have $\bar\nu^2_{\bar\tau_{en}} = 0$ in (107) by (91). For a nonreducible touch point $\bar\tau_{to}$ introduced as a touch point in $\mathcal{F}_j$, (101) becomes

$$0 = g_y^{(1)}(\bar y(\bar\tau_{to}))z(\bar\tau_{to}). \tag{109}$$

The above constraint holds as well for nonreducible touch points converted into boundary arcs by (97). Moreover, we substitute $\sigma_{\bar\tau_{to}}$ using (101) into the jump condition (108) for reducible touch points, and we consider for nonreducible touch points introduced as touch points the constraint (109) with associated multiplier $\sigma_{\bar\tau_{to}}\bar\nu_{\bar\tau_{to}}$ in (108). In this way we obtain that (95)–(97) and (100)–(108) constitute the first-order optimality condition of the linear-quadratic problem (PQ) of minimizing

$$\begin{aligned}
\mathcal{Q}^2(v) \;:=\; & \int_0^T D^2_{(u,y)(u,y)}\tilde H(\bar u,\bar y,\bar p_2,\bar\eta_2)((v,z_v),(v,z_v))\mathrm{d}t + \phi_{yy}(\bar y(T))(z_v(T),z_v(T)) \\
+ & \sum_{\bar\tau_{en}\in\bar{\mathcal{T}}_{en}} \Big(\nu^1_{\bar\tau_{en}} g_{yy}(\bar y(\bar\tau_{en}))(z_v(\bar\tau_{en}),z_v(\bar\tau_{en})) + \nu^2_{\bar\tau_{en}} g^{(1)}_{yy}(\bar y(\bar\tau_{en}))(z_v(\bar\tau_{en}),z_v(\bar\tau_{en}))\Big) \\
+ & \sum_{\bar\tau_{to}\in\bar{\mathcal{T}}_{to}} \bar\nu_{\bar\tau_{to}} g_{yy}(\bar y(\bar\tau_{to}))(z_v(\bar\tau_{to}),z_v(\bar\tau_{to})) \;-\; \sum_{\bar\tau_{to}\in\bar{\mathcal{T}}^{ess}_{red}} \bar\nu_{\bar\tau_{to}} \frac{(g_y^{(1)}(\bar y(\bar\tau_{to}))z_v(\bar\tau_{to}))^2}{g^{(2)}(\bar u(\bar\tau_{to}),\bar y(\bar\tau_{to}))},
\end{aligned}$$

subject to the constraints (96), (97), (100), (105), and (109) at nonreducible touch points. Since $\frac{\mathrm{d}}{\mathrm{d}t}g_y(\bar y(t))z_v(t) = g_y^{(1)}(\bar y)z_v$ and $\frac{\mathrm{d}^2}{\mathrm{d}t^2}g_y(\bar y(t))z_v(t) = g_y^{(2)}(\bar u,\bar y)z_v + g_u^{(2)}(\bar u,\bar y)v$, the constraints (96), (97), and (105) are equivalent to $g_y(\bar y(t))z(t) = 0$ on boundary arcs (of positive length) $[\bar\tau_{en},\bar\tau_{ex}]$. Consequently, the constraints (96), (97), (100), (105), and (109) of (PQ) are equivalent to (63), (92), and $g_y(\bar y(\bar\tau_{to}))z(\bar\tau_{to}) = 0$ for all nonessential touch point $\bar\tau_{to}$ introduced in the shooting mapping $\mathcal{F}_j$.

By straightforward calculation (see [8, Lemma 3.6] and [16, Lemma 3.1]), we can show that the quadratic form $\mathcal{Q}^2(v)$ is equal to the left-hand side of (93). Since the latter is a Legendre form by assumption (20), (93) implies that (PQ) has a weakly lower semi-continuous and strongly convex cost function on its closed and convex feasible set. Moreover, the constraints of (PQ) are onto by assumption (A2) (see Lemma 2.5) and hence the unique solution and associated multipliers of the first-order optimality condition of (PQ) are zero. This implies that $(v,z,\pi_2,\zeta_2)\equiv 0$. Therefore, $\pi_0 = 0$ and the multipliers associated with the constraints (96)–(97), (100), and (109) for nonreducible touch points introduced as touch points are equal to zero, implying that

$$\gamma^1_{\bar\tau_{en}} = 0, \qquad \gamma^2_{\bar\tau_{en}} + \sigma_{\bar\tau_{en}}\bar\nu^1_{\bar\tau_{en}} = 0, \qquad \gamma_{\bar\tau_{to}} = 0, \tag{110}$$

and, for nonreducible touch points $\bar\tau_{to}$ introduced as touch points,

$$\sigma_{\bar\tau_{to}}\bar\nu_{\bar\tau_{to}} = 0. \tag{111}$$

By (98)–(99), since $\frac{\mathrm{d}}{\mathrm{d}t}g^{(2)}(\bar u(t),\bar y(t))|_{t=\bar\tau^-_{en},\bar\tau^+_{ex}} \neq 0$ both for entry/exit points of boundary arcs by Lemma 2.3, and for nonreducible touch points converted into boundary arcs by hypothesis (A6)(i), we have that $\sigma_{\bar\tau_{en}} = 0 = \sigma_{\bar\tau_{ex}}$, and by (101), $\sigma_{\bar\tau_{to}} = 0$ for reducible touch points $\bar\tau_{to}$. Finally, with (110)–(111), since $\bar\nu_{\bar\tau_{to}} \neq 0$ at nonreducible touch points $\bar\tau_{to}$ by (A6)(i) and Lemma 4.3, it follows that $\omega = 0$, i.e. the Jacobian of the shooting mapping $\mathcal{F}_j$ is one-to-one, and hence invertible. □

6.2 Stability of shooting parameters

Let $(\bar u,\bar y)$ be a stationary point of $(\mathcal{P})$ satisfying (A2)–(A6) and the strong second-order sufficient condition (65) and let $(\mathcal{P}^\mu)$ be a stable extension of $(\mathcal{P})$. For μ in a neighborhood of $\bar\mu$, the perturbed problem $(\mathcal{P}^\mu)$ has by Th. 5.1 a locally unique stationary point (u^μ,y^μ), which has by Theorems 3.1 and 4.4 a neighboring structure to that of $(\bar u,\bar y)$, in the sense of Def. 6.5. Therefore it makes sense to speak about the shooting parameters associated with (u^μ,y^μ). Note that its set of shooting parameters may not necessarily be unique if (u^μ,y^μ) has nonessential or nonreducible touch points, since a nonessential touch point may or not be introduced in the set of shooting parameters, with an associated zero jump parameter, and a nonreducible touch point may be considered either as a boundary arc (of zero length) or as a touch point. The next lemma shows that the stationary point (u^μ, y^μ) of $(\mathcal{P}^\mu)$ has its shooting parameters in the neighborhood of the shooting parameters of $(\bar{u}, \bar{y})$.

For this it will be useful to make explicit the relation between the multipliers η_2 and η^2 used respectively in the shooting algorithm and in the stability analysis. Recall that the multipliers used in the shooting algorithm are defined by (69)–(70) while those used in the stability analysis are defined by (11)–(12). Moreover, by [8, Prop. 2.10], for all boundary arcs (τ_{en}, τ_{ex}) (including the case $\tau_{en} = \tau_{ex}$), we have that

$$\nu^1_{\tau_{en}} = \int_{[\tau_{en}, \tau_{ex}]} \mathrm{d}\eta = [\eta(\tau_{en})] + \eta_1(\tau_{en}^+), \tag{112}$$

and the condition (81) holds *a fortiori* for a stationary point. Combining the above relations, we obtain that

$$\begin{aligned} \eta^1(t) &= \eta_1(t) + \sum_{\tau_{en} \in \mathcal{T}_{en}} \nu^1_{\tau_{en}} \mathbf{1}_{[0,\tau_{en}]}(t) + \sum_{\tau_{to} \in \mathcal{T}_{to}} \nu_{\tau_{to}} \mathbf{1}_{[0,\tau_{to}]}(t), & (113) \\ \eta^2(t) &= \int_t^T \eta^1(s)\mathrm{d}s \\ &= \eta_2(t) + \sum_{\tau_{en} \in \mathcal{T}_{en}} \mathbf{1}_{[0,\tau_{en}]}(t)(\nu^2_{\tau_{en}} + \nu^1_{\tau_{en}}(\tau_{en} - t)) + \sum_{\tau_{to} \in \mathcal{T}_{to}} \nu_{\tau_{to}} \mathbf{1}_{[0,\tau_{to}]}(t)(\tau_{to} - t). & (114) \end{aligned}$$

Here $\mathbf{1}_{[a,b]}(\cdot)$ denotes the indicator function of the interval $[a,b] \subset [0,T]$ equal to 1 on $[a,b]$ and zero outside. Then p_2 and p^2 defined respectively by (12) and (70) are related by

$$p^2 \ = \ p_2 - (\eta^1 - \eta_1)g_y(y) - (\eta^2 - \eta_2)g_y^{(1)}(y). \tag{115}$$

Lemma 6.7. *Let $(\bar{u}, \bar{y})$ be a stationary point of $(\mathcal{P})$ satisfying (A2)–(A6) and the strong second-order sufficient condition* (65) *and let $(\mathcal{P}^\mu)$ be a stable extension of $(\mathcal{P})$. Then for each $\varepsilon > 0$, there exist neighborhoods W of $\bar{\mu}$ and V_∞ of $\bar{u}$ (in L^∞) such that for each $\mu \in W$, the locally unique stationary point (u^μ, y^μ) of $(\mathcal{P}^\mu)$ with $u^\mu \in V_\infty$ has a neighboring structure to that of $(\bar{u}, \bar{y})$. Moreover, any vector of shooting parameters θ^μ associated with (u^μ, y^μ), of appropriate dimension, satisfies*

$$|\theta^\mu - \bar{\theta}_j| < \varepsilon$$

where $\bar{\theta}_j$ is the vector of shooting parameters associated with $(\bar{u}, \bar{y})$ matching the structure of θ^μ.

It follows from the above lemma and Lemma 6.6 that for a given neighboring structure $\mathcal{F}_j$ of $(\bar{u}, \bar{y})$, the vector of shooting parameters θ^μ associated with the stationary point (u^μ, y^μ) is locally unique.

Proof. It follows from Theorems 3.1 and 4.4 that the locally unique (by Th. 5.1) stationary point (u^μ, y^μ) of $(\mathcal{P}^\mu)$ has a neighboring structure to that of $(\bar{u}, \bar{y})$. The convergence of junction times was proved in Theorems 3.1 and 4.4. So let us show the convergence of jump parameters. For this we use the formula (114) that links the multiplier η^2 used in the stability analysis to the shootings parameters and the uniform convergence of $\eta^{2,\mu}$ towards $\bar{\eta}^2$ by Prop. 2.6(iii). The proof is by finite induction.

Let N denote the total number of boundary arcs and touch points of the trajectory $(\bar{u}, \bar{y})$. We may write that $\{1, \dots, N\} = \mathcal{N}_{ba} \cup \mathcal{N}_{to}$, where $\mathcal{N}_{ba} \cap \mathcal{N}_{to} = \emptyset$ and $\mathcal{N}_{ba}$ and $\mathcal{N}_{to}$ denote the sets of index corresponding respectively to boundary arcs (possibly of zero length) and to touch points (possibly nonreducible or nonessential). This partition is not unique since a nonreducible touch point can be considered either as a boundary arc of zero length or as a touch point. We have then that $I(g(\bar{y})) = \cup_{i=1}^N \bar{I}_i$, where $\bar{I}_i := [\bar{\tau}_{en,i}, \bar{\tau}_{ex,i}]$ for $i \in \mathcal{N}_{ba}$ (with possibly $\bar{\tau}_{en,i} = \bar{\tau}_{ex,i}$), $\bar{I}_i := \{\bar{\tau}_{to,i}\}$ for $i \in \mathcal{N}_{to}$, $\bar{I}_i \cap \bar{I}_j = \emptyset$ for $i \neq j$, and $\bar{I}_i < \bar{I}_{i+1}$ for all $i < N$ (in the sense that $t < t'$ for all $(t,t') \in \bar{I}_i \times \bar{I}_{i+1}$). The jump parameters associated with a boundary arc $[\bar{\tau}_{en,i}, \bar{\tau}_{ex,i}]$ are denoted by $\bar{\nu}_i^1$ and $\bar{\nu}_i^2$ and that associated with a touch point $\bar{\tau}_{to,i}$ by $\bar{\nu}_i$. Since (u^μ, y^μ) has a neighboring structure to that of $(\bar{u}, \bar{y})$, we can choose the partition $(\mathcal{N}_{ba}, \mathcal{N}_{to})$ such that $I(g^\mu(y^\mu)) = \cup_{i=1}^N I_i^\mu$ for a sequence $\mu_n \to_{n\to\infty} \bar{\mu}$, where $I_i^\mu = [\tau^\mu_{en,i}, \tau^\mu_{ex,i}]$ for $i \in \mathcal{N}_{ba}$, with associated jump parameters $\nu_i^{1,\mu}$ and

$\nu_i^{2,\mu}$, and $I_i^\mu = \{\tau_{to,i}^\mu\}$ with jump parameter ν_i^μ or possibly $I_i^\mu = \emptyset$ (if $\bar\tau_{to,i}$ is a nonessential touch point) for $i \in \mathcal{N}_{to}$.

Given $k \in \{1,\ldots,N\}$, assume by induction that the jump parameters associated with $I_i^{\mu_n}$ converge to those associated with $\bar I_i$ for all $i \in \{k+1,\ldots,N\}$. (For $k = N$ we assume nothing.) Let us show that the jump parameters associated with $I_k^{\mu_n}$ converges to those associated with $\bar I_k$. There are two cases to consider.

Case 1: $k \in \mathcal{N}_{to}$. If $I_k^{\mu_n} = \emptyset$, there is nothing to prove, so assume that $I_k^{\mu_n} = \{\tau_{to,k}^{\mu_n}\}$. Recall that by definition, η_1 and η_2 vanish on interior arcs. Then for a fixed $\varepsilon > 0$ small enough ($\varepsilon < \min_{\tau,\tau'\in\bar{\mathcal{T}}\cup\{0\},\tau\neq\tau'} \frac{1}{2}|\tau-\tau'|$), for all $t \in [\bar\tau_{to,k} - 2\varepsilon, \bar\tau_{to,k} - \varepsilon]$, we have by (114) and Th. 5.1 for n large enough that

$$\begin{aligned}\eta^{2,\mu_n}(t) &= \sum_{i\in\mathcal{N}_{ba},i>k} (\nu_i^{2,\mu_n} + \nu_i^{1,\mu_n}(\tau_{en,i}^{\mu_n} - t)) + \sum_{i\in\mathcal{N}_{to},i>k} \nu_i^{\mu_n}(\tau_{to,i}^{\mu_n} - t) \ + \ \nu_k^{\mu_n}(\tau_{to,k}^{\mu_n} - t)\\ &\underset{n\to\infty}{\longrightarrow} \ \bar\eta^2(t) = \sum_{i\in\mathcal{N}_{ba},i>k} (\bar\nu_i^2 + \bar\nu_i^1(\bar\tau_{en,i} - t)) + \sum_{i\in\mathcal{N}_{to},i>k} \bar\nu_i(\bar\tau_{to,i} - t) \ + \ \bar\nu_k(\bar\tau_{to,k} - t).\end{aligned}$$

Since the junction times of (u^{μ_n}, y^{μ_n}) converge to those of $(\bar u, \bar y)$, as well as the jump parameters associated with $I_i^{\mu_n}$ for $i > k$ by the induction hypothesis, we deduce immediately that $\nu_k^{\mu_n}$ converges to $\bar\nu_k$.

Case 2: $k \in \mathcal{N}_{ba}$. Then $I_k^{\mu_n} = [\tau_{en,k}^{\mu_n}, \tau_{ex,k}^{\mu_n}]$ and reasoning similarly, for a fixed $\varepsilon > 0$ small enough, for all $t \in [\bar\tau_{en,k} - 2\varepsilon, \min\{\bar\tau_{en,k}, \tau_{en,k}^{\mu_n}\})$ and n large enough, we have that

$$\begin{aligned}\eta^{2,\mu_n}(t) &= \sum_{i\in\mathcal{N}_{ba},i>k} (\nu_i^{2,\mu_n} + \nu_i^{1,\mu_n}(\tau_{en,i}^{\mu_n} - t)) + \sum_{i\in\mathcal{N}_{to},i>k} \nu_i^{\mu_n}(\tau_{to,i}^{\mu_n} - t)\\ &\qquad + \nu_k^{2,\mu_n} + \nu_k^{1,\mu_n}(\tau_{en,k}^{\mu_n} - t)\\ &\underset{n\to\infty}{\longrightarrow} \ \bar\eta^2(t) = \sum_{i\in\mathcal{N}_{ba},i>k} (\bar\nu_i^2 + \bar\nu_i^1(\bar\tau_{en,i} - t)) + \sum_{i\in\mathcal{N}_{to},i>k} \bar\nu_i(\bar\tau_{to,i} - t)\\ &\qquad + \bar\nu_k^2 + \bar\nu_k^1(\bar\tau_{en,k} - t).\end{aligned}$$

By letting $t \uparrow \bar\tau_{en,k}$, $t < \min\{\bar\tau_{en,k}, \tau_{en,k}^{\mu_n}\}$, and $n \to +\infty$, using that the convergence of η^{2,μ_n} towards $\bar\eta^2$ is uniform and that ν_k^{1,μ_n} is bounded (since $\nu_k^{1,\mu_n} = \int_{[\tau_{en,k}^{\mu_n}, \tau_{ex,k}^{\mu_n}]} d\eta^{\mu_n}$ by (112) and $d\eta^{\mu_n}$ is uniformly bounded by Prop. 2.6(i)), we deduce as previously that $\nu_k^{2,\mu_n} \to \bar\nu_k^2$. Taking then $t \in [\bar\tau_{en,k} - 2\varepsilon, \bar\tau_{en,k} - \varepsilon]$, it follows that $\nu_k^{1,\mu_n} \to \bar\nu_k^1$. This completes the induction step and achieves to show the converge of jump parameters of (u^μ, y^μ) towards those of $(\bar u, \bar y)$.

It remains to show the convergence of the initial costate. For a small $\varepsilon > 0$ and $\|\mu - \bar\mu\|$ small enough, the state constraint $g^\mu(y^\mu)$ is not active on $[0, \varepsilon]$. Therefore, by (113), $\eta^{1,\mu}$ converges uniformly to $\bar\eta^1$ on $[0, \varepsilon]$ and $\eta_1^\mu = \eta_2^\mu = 0$ since we are on an interior arc. It follows that for all $t \in [0, \varepsilon]$, using (115),

$$\begin{aligned}p_2^\mu(t) &= p^2(t) + \eta^{1,\mu}(t) g_y^\mu(y^\mu(t)) + \eta^{2,\mu}(t)(g^\mu)_y^{(1)}(y^\mu(t))\\ &\underset{\mu\to\bar\mu}{\longrightarrow} \ \bar p^2(t) + \bar\eta^1(t) g_y(\bar y(t)) + \bar\eta^2(t) g_y^{(1)}(\bar y(t)) \ = \ \bar p_2(t)\end{aligned}$$

since $p^{2,\mu}$, $\eta^{2,\mu}$ and y^μ converges uniformly to $\bar p^2$, $\bar\eta^2$ and $\bar y$, respectively. For $t = 0$ this gives the convergence of the initial costate $p_2^\mu(0) \to \bar p_2(0)$. This achieves the proof of the lemma. □

6.3 Additional conditions for a stationary point

By Lemma 6.7, we know that the locally unique stationary point (u^μ, y^μ) of the perturbed problem $(\mathcal{P}^\mu)$ has its shooting parameters in the neighborhood of those of the reference trajectory $(\bar u, \bar y)$. By Lemma 6.6, the shooting algorithm is then well-posed to find a vector of shooting parameters associated with (u^μ, y^μ). Lemma 6.7 ensures that *at least one* of the solutions θ_j^μ obtained in Lemma 6.6 for the neighboring structures to that of $(\bar u, \bar y)$ is associated to this (locally unique) stationary point of $(\mathcal{P}^\mu)$. Of course we do not know a priori what the structure of (u^μ, y^μ) is. We only know that it is a neighboring structure to that of $(\bar u, \bar y)$. In Lemma 6.9 below we give necessary and sufficient conditions in order to recognize a vector of shooting parameters associated with a stationary point of the perturbed problem, among all the solutions of (94). Let us first note the following.

Remark 6.8. The statement of Lemma 6.1 extends without difficulty to the case when there are nonreducible touch points converted into boundary arcs of zero length (with $\tau_{en} = \tau_{ex}$). In that case $\eta_2(\tau_{en}^+) = 0$ and (81) yields that $\nu^2_{\tau_{en}} = 0$ automatically holds at nonreducible touch points converted into boundary arcs. The statement of Prop. 6.2 extend as well. For nonreducible touch points τ_{to} converted into boundary arcs of zero length, since $\dot\eta_2(\tau_{en}^+) = 0$, (84) amounts to the classical condition $\nu_{\tau_{to}} = \nu^1_{\tau_{en}} \geq 0$, while (85) is automatically satisfied (with equality).

Lemma 6.9. *Let $(\bar u, \bar y)$ be a stationary point of $(\mathcal{P})$ satisfying (A2)–(A6) and (93). For $j \in \{1, \dots, N_s\}$, let $\mathcal{F}_j$ denote one of the shooting mappings associated with a neighboring structure to $(\bar u, \bar y)$. Then there exist a neighborhood W of $\bar\mu$ and a neighborhood V_j of $\bar\theta_j$, such that a solution θ in V_j of (94) for $\mu \in W$ is associated with a stationary point of $(\mathcal{P}^\mu)$ iff, denoting by (u, y, p_2, η_2) the trajectory and multipliers associated with θ, the following conditions are satisfied:*

$$
\begin{aligned}
0 &\geq g^\mu(y(t)) \text{ on } [0,T], && (116)\\
0 &\geq (g^\mu)^{(2)}(u(\tau_{to}), y(\tau_{to})) \quad \text{for each touch point } \tau_{to} \text{ of } \theta, && (117)\\
0 &\leq \nu_{\tau_{to}} \quad \text{for each touch point } \tau_{to} \text{ of } \theta, && (118)\\
\tau_{en} &\leq \tau_{ex} \quad \text{for each boundary arc of } \theta. && (119)
\end{aligned}
$$

Proof. By Prop. 6.2, it is obvious that the conditions (116), (118), and (119) are necessary for a stationary point. The condition (117) is necessary as well, since in the neighborhood of a touch point τ, we have that

$$g^\mu(y(t)) = (g^\mu)^{(2)}(u(\tau), y(\tau))\frac{(t-\tau)^2}{2} + o(|t-\tau|^2) \leq 0.$$

Now we show that the conditions (116)–(119) are sufficient to have a stationary point of $(\mathcal{P}^\mu)$. In order to show that the trajectory and multipliers (u, y, p_2, η_2) associated with θ are a stationary point of $(\mathcal{P}^\mu)$ and its associated multipliers in the shooting algorithm, we have to show by Prop. 6.2 and Remark 6.8 that the additional conditions (82)–(86) are satisfied.

The conditions (82) and (86) at touch points follow immediately from (116) and (118). Let us show now (84)–(85). By (A2) and (A3), implying (78), for μ in the vicinity of $\bar\mu$, (u, y, p_2, η_2) satisfies by continuity (79)–(80) and hence, it follows from Lemma 6.1 and Remark 6.8 that u is continuous over $[0,T]$. Therefore the conditions (84)–(85) at entry and exit points of boundary arcs of nonzero length (τ_{en}, τ_{ex}) are satisfied by Rem. 6.4 as a consequence of (116). For possible boundary arcs of zero length $\tau_{en} = \tau_{ex}$, (84)–(85) amounts to check that $\nu^1_{\tau_{en}} \geq 0$. By the same arguments than in the proof of Lemma 6.3, this last condition is equivalent to $[g^{(3)}(\dot u(\tau_{en}), u(\tau_{en}), y(\tau_{en}))] \leq 0$, which holds by continuity for $\|\mu - \bar\mu\|$ and $|\theta - \bar\theta_j|$ small enough by (A6)(i).

Let us end the proof by showing that (83) is satisfied on boundary arcs (τ_{en}, τ_{ex}) with $\tau_{en} < \tau_{ex}$. *Define* the multipliers η^2 and p^2 by respectively (114) and (115). By (81), we have that η^2 is continuous over $[0,T]$. By (113)–(115) and (76)–(77), we see directly that p^2 is continuous over $[0,T]$ as well. Moreover, (72)–(73) imply by straightforward calculations that the following hold over $[0,T]$

$$
\begin{aligned}
-\dot p^2 &= \tilde H^\mu_y(u, y, p^2, \eta^2), && (120)\\
0 &= \tilde H^\mu_u(u, y, p^2, \eta^2). && (121)
\end{aligned}
$$

On the interior of each arc, (u, η_2) can be expressed as a C^1 function of (y, p_2) and μ. Therefore, for $\|\mu - \bar\mu\|$ and $|\theta - \bar\theta_j|$ small enough, we have that $|u(t) - \bar u(t)|$, $|y(t) - \bar y(t)|$, $|\eta_2(t) - \bar\eta_2(t)|$, $|p_2(t) - \bar p_2(t)|$ are arbitrarily small, uniformly on an interior of each arc. Since u, y, η^2, and p^2 are continuous, uniformly over $[0,T]$, we deduce that $\|u - \bar u\|_\infty, \|y - \bar y\|_\infty, \|\eta^2 - \bar\eta^2\|_\infty, \|p^2 - \bar p^2\|_\infty$ are arbitrarily small for μ and θ in the neighborhood of $\bar\mu$ and $\bar\theta_j$, respectively. Using the relations (120)–(121), we obtain like in section 3 that the relation (40) holds. From now, the end of the proof is similar to the end of the proof of Th. 3.1 or 4.4 to show that the uniform strict complementarity assumption holds on boundary arc, depending on whether (τ_{en}, τ_{ex}) is in the neighborhood of a boundary arc $(\bar\tau_{en}, \bar\tau_{ex})$ or in the neighborhood of a nonreducible touch point $\bar\tau_{to}$ of $(\bar u, \bar y)$. ∎

7 Application to homotopy methods

In this section, we extend to second-order state constraints the homotopy algorithm of [7] that detects automatically the structure of the trajectory for first-order state constraints, in the case when assumptions (A2)–(A6) and the strong second-order sufficient condition (65) are satisfied along the homotopy path.

7.1 Description of the algorithm

We consider the natural homotopy on the state constraint, for $\mu \in [0,1]$,

$$g^{\mu}(y) \ := \ g(y) - (1-\mu)M \qquad \text{and} \qquad (\ell^{\mu}, \phi^{\mu}, f^{\mu}, y_0^{\mu}) \equiv (\ell, \phi, f, y_0), \tag{122}$$

where $M > 0$ is large enough, so that the state constraint of problem $(\mathcal{P}^0)$ is not active, and we have that $(\mathcal{P}^1) \equiv (\mathcal{P})$. More generally, the algorithm below can be extended to any stable extension $(\mathcal{P}^{\mu})$ of $(\mathcal{P})$ satisfying the assumption (H0) below, if a solution of $(\mathcal{P}^0)$ can be easily obtained:

(H0) $(\mathcal{P}^{\mu})$ is a stable extension of $(\mathcal{P})$, defined for $\mu \in [0,1]$, such that $(\mathcal{P}^1) \equiv (\mathcal{P})$ and satisfying $g^{\mu}(y_0^{\mu}) < 0$ for all $\mu \in [0,1]$.

The homotopy algorithm is as follows. We denote the current structure of the trajectory by $\mathcal{S}$, i.e. the variable $\mathcal{S}$ indicates the number and order of boundary arcs and touch points. The shooting mapping associated with the structure $\mathcal{S}$ is denoted by $\mathcal{F}_{\mathcal{S}}$. Given a vector of shooting parameters θ, of dimension appropriate with $\mathcal{S}$, and a value $\mu \in [0,1]$ of the homotopy parameter, we will denote by $(u^{\mu}_{\mathcal{S},\theta}, y^{\mu}_{\mathcal{S},\theta})$ the trajectory associated with θ in the shooting algorithm for the structure $\mathcal{S}$ and the homotopy parameter μ.

Algorithm 7.1 (Homotopy Algorithm)**.**

Input p_0 *initial costate candidate for the unconstrained problem* $(\mathcal{P}^0)$ *and* $\delta \in (0,1)$.

Initialization *Let $\mathcal{S}$ be the empty structure (with no boundary arc and no touch point). Solve by the Newton algorithm (initialized by the value p_0) $\mathcal{F}_{\mathcal{S}}(\theta, 0) = 0$ and obtain a vector of shooting parameters θ associated with a solution of the unconstrained problem $(\mathcal{P}^0)$. Set $M := \max_{t\in[0,T]} g(y^1_{\mathcal{S},\theta}(t))$. If $M \le 0$ then $\mu := 1$ else $\mu := 0$. Set $\Delta\mu := \delta$.*

While $\mu < 1$ **do**

Prediction Step *Set $\bar{\mu} := \min\{\mu + \Delta\mu; 1\}$ and compute*

$$\bar{\theta} := \theta - D_{\theta}\mathcal{F}_{\mathcal{S}}(\theta,\mu)^{-1} D_{\mu}\mathcal{F}_{\mathcal{S}}(\theta,\mu)\Delta\mu. \tag{123}$$

Correction Step *Solve, with the Newton algorithm initialized by the value $\bar{\theta}$,*

$$\mathcal{F}_{\mathcal{S}}(\hat{\theta}, \bar{\mu}) = 0. \tag{124}$$

If *the Newton algorithm fails, set $\Delta\mu := \Delta\mu/2$ and* **go to** *the* prediction step*;*
Else *obtain a vector of shooting parameters $\hat{\theta}$ solution of* (124)*.*

Update the structure

[to→ba] **If** *there exists a touch point τ_{to} of $\hat{\theta}$ such that*

$$(g^{\bar{\mu}})^{(2)}(u^{\bar{\mu}}_{\mathcal{S},\hat{\theta}}(\tau_{to}), y^{\bar{\mu}}_{\mathcal{S},\hat{\theta}}(\tau_{to})) \ge 0, \tag{125}$$

let $\hat{\mathcal{S}}$ be the structure obtained by replacing in $\mathcal{S}$ the touch point τ_{to} by a boundary arc, set $\mathcal{S} := \hat{\mathcal{S}}$, and let $\bar{\theta}$ be the vector of shooting parameters obtained from θ by replacing the touch point τ_{to} and its jump parameter $\nu_{\tau_{to}}$ by a boundary arc, with shooting parameters

$$\tau_{en} := \tau_{to}, \qquad \tau_{ex} := \tau_{to}, \qquad \nu^1_{\tau_{en}} := \nu_{\tau_{to}}, \qquad \nu^2_{\tau_{en}} := 0. \tag{126}$$

Go to *the* correction step*;*

[ADD TO] **Else if** $m := \max_{t\in[0,T]} g^{\bar{\mu}}(y^{\bar{\mu}}_{\mathcal{S},\hat{\theta}}(t)) > 0$, *set* $\tau_{to} := \operatorname{argmax}_{t\in[0,T]} g^{\bar{\mu}}(y^{\bar{\mu}}_{\mathcal{S},\hat{\theta}}(t))$, *let* $\hat{\mathcal{S}}$ *be the structure obtained from* $\mathcal{S}$ *by adding the touch point* τ_{to}, *set* $\mathcal{S} := \hat{\mathcal{S}}$, *and let* $\bar{\theta}$ *be the vector of shooting parameters obtained from* θ *by adding the touch point* τ_{to} *with a zero jump parameter* $\nu_{\tau_{to}}$. **Go to** *the* CORRECTION STEP*;*

[REM TO] **Else if** *there exists a touch point* τ_{to} *of* $\hat{\theta}$ *such that its jump parameter* $\nu_{\tau_{to}}$ *is negative, then let* $\hat{\mathcal{S}}$ *be the structure obtained from* $\mathcal{S}$ *by deleting the touch point* τ_{to}, *set* $\mathcal{S} := \hat{\mathcal{S}}$, *and let* $\bar{\theta}$ *be the vector of shooting parameters obtained from* θ *by deleting the touch point* τ_{to} *and its jump parameter* $\nu_{\tau_{to}}$. **Go to** *the* CORRECTION STEP*;*

[BA→TO] **Else if** *there exists a boundary arc* (τ_{en}, τ_{ex}) *of* $\hat{\theta}$ *such that* $\tau_{en} > \tau_{ex}$, *then let* $\hat{\mathcal{S}}$ *be the structure obtained from* $\mathcal{S}$ *by replacing the boundary arc* (τ_{en}, τ_{ex}) *by a touch point, set* $\mathcal{S} := \hat{\mathcal{S}}$, *and let* $\bar{\theta}$ *be the vector of shooting parameters obtained from* θ *by replacing the shooting parameters associated with the boundary arc* (τ_{en}, τ_{ex}) *by a touch point and its jump parameter,*

$$\tau_{to} := \tau_{en}, \qquad \nu_{\tau_{to}} := \nu^1_{\tau_{en}}. \tag{127}$$

Go to *the* CORRECTION STEP*;*

[OK] **Else** *set* $\theta := \hat{\theta}$, $\mu := \bar{\mu}$.

End While

7.2 Construction of the homotopy path

The analysis of the existence of the homotopy path is analogous to that of [7] for first-order state constraints. Let $(\mathcal{P}^\mu)$ satisfy (H0) and assume that

(H1) The problem $(\mathcal{P}^0)$ has a local solution (u^0, y^0) satisfying (A2)–(A6) and the strong second-order sufficient condition (65).

By Th. 5.1, there exists $\delta > 0$ such that for all $\mu \in [0, \delta)$, $(\mathcal{P}^\mu)$ has a stationary point (u^μ, y^μ), locally unique in a L^∞-neighborhood of (u^0, y^0), which is Hölder continuous w.r.t. μ in the L^∞ norm and is a local solution of $(\mathcal{P}^\mu)$. By assumptions (A4)–(A6) and Theorems 3.1 and 4.4, (u^μ, y^μ) has a neighboring structure to that of (u^0, y^0) (in the sense of Def. 6.5), i.e. satisfies (A4). Moreover, reducing δ if necessary, assumptions (A2)–(A3) are satisfied, as well as (A5) by Theorems 3.1 and 4.4 and (A6) by continuity. Finally, by Th. 5.1, (u^μ, y^μ) satisfies the strong second-order sufficient condition (65). So let

$$\begin{aligned}\mu_{max} := \sup\{\mu \in [0,1] : &\text{for all } \mu' \in [0,\mu], \text{ the locally unique solution } (u^{\mu'}, y^{\mu'}) \text{ of } (\mathcal{P}^{\mu'}) \\ &\text{satisfy (A2)–(A6) and the strong second-order sufficient condition (65).}\}\end{aligned}$$

Under assumption (H1), we have that $\mu_{max} > 0$. We obtain the following result.

Lemma 7.2. *Assume that (H0)–(H1) are satisfied, and that there exist* $L, \beta, \sigma > 0$ *such that for all* $\mu \in [0, \mu_{max})$,

$$\|u^\mu\|_{1,1} \leq L, \tag{128}$$

$$|(g^\mu)^{(2)}_u(u^\mu(t), y^\mu(t))| \geq \beta \quad \text{for all } t \in I_\sigma(g^\mu(y^\mu)). \tag{129}$$

Then for all sequences $\mu_n \uparrow \mu_{max}$, *there exists a subsequence, still denoted by* (μ_n), *such that* $(u^{\mu_n}, y^{\mu_n}, p^{2,\mu_n}, \eta^{2,\mu_n})$ *converges uniformly to some* $(\tilde{u}, \tilde{y}, \tilde{p}^2, \tilde{\eta}^2)$, *and* $(\tilde{u}, \tilde{y}, \tilde{p}^2, \tilde{\eta}^2)$ *is a stationary point and its alternative multipliers of* $(\mathcal{P}^{\mu_{max}})$.

Moreover, if $(\tilde{u}, \tilde{y}, \tilde{p}^2, \tilde{\eta}^2)$ *satisfies assumptions (A2)–(A6) and the strong second-order sufficient condition (65), then* $(u^\mu, y^\mu, p^{2,\mu}, \eta^{2,\mu})$ *converges uniformly when* $\mu \uparrow \mu_{max}$ *to a locally unique local solution of* $(\mathcal{P}^{\mu_{max}})$ *and its alternative multipliers* $(\tilde{u}, \tilde{y}, \tilde{p}^2, \tilde{\eta}^2) =: (u^{\mu_{max}}, y^{\mu_{max}}, p^{2,\mu_{max}}, \eta^{2,\mu_{max}})$, *and* $\mu_{max} = 1$, *i.e. the homotopy path is well-defined over* $\mu \in [0, 1]$.

Proof. The proof follows from that of [7, Lemma 8.4]. By the compactness Theorem in BV [2, Th. 3.23], the weak-* convergence in $\mathcal{M}[0,T]$ of the multiplier $d\eta^{\mu_n}$ associated with (u^{μ_n}, y^{μ_n}) in the optimality conditions (7)–(10) implies the uniform convergence of the alternative multiplier η^{2,μ_n} defined by (11). The uniform convergence of p^{2,μ_n} follows then from (16). □

Given a stable extension $(\mathcal{P}^\mu)$ satisfying (H0) and (H1), we make the following assumptions that guarantee the existence (and local uniqueness) of the homotopy path over $\mu \in [0,1]$:

(H2) There exists $L, \beta, \sigma > 0$ such that for all $\mu \in [0,1]$, (u^μ, y^μ) satisfies (128)–(129);

(H3) For all $\mu \in [0,1]$, (u^μ, y^μ) satisfies the assumptions (A3)–(A6);

(H4) For all $\mu \in [0,1]$, (u^μ, y^μ) satisfies the strong second-order sufficient condition (65).

7.3 Proof of convergence

In addition to hypotheses (H0)–(H4), we make the assumptions below in the proof of correctness of Algorithm 7.1. Note that a change in the structure of the trajectories (u^μ, y^μ), $\mu \in [0,1]$, may occur only at some values $\tilde{\mu} \in [0,1)$ having either a nonessential or a nonreducible touch point.

(H5) There exist finitely values of μ, $0 \leq \tilde{\mu}_1 < \ldots < \tilde{\mu}_N < 1$ for which the structure of the trajectory changes.

(H6) For each $\tilde{\mu}_k$, $k = 1, \ldots, N$, $(u^{\tilde{\mu}_k}, y^{\tilde{\mu}_k})$ has either one (single) nonessential touch point or one (single) nonreducible touch point.

When (H6) holds, there are only two different neighboring structures to that of $(u^{\tilde{\mu}_k}, y^{\tilde{\mu}_k})$, for each $\tilde{\mu}_k$. The algorithm 7.1 could be generalized to the case when (H6) does not hold, but in that case the Update the structure step is more delicate. A possibility is to enumerate all the possible neighboring structures until the conditions (116)–(119) of Lemma 6.9 are satisfied.

Proposition 7.3. *Let $(\mathcal{P}^\mu)$ be given by* (122) *and assume that assumptions (H1)–(H6) are satisfied. Then there exist a neighborhood V_0 of $p_2^0(0)$, the initial costate of the unconstrained problem $(\mathcal{P}^0)$, and $\bar{\delta} > 0$ such that for all $p_0 \in V_0$ and all $\delta \in (0, \bar{\delta})$, the homotopy algorithm 7.1 follows the homotopy path and ends with a vector of shooting parameters θ, of appropriate dimension, associated with a local solution (u^1, y^1) of $(\mathcal{P}^1) \equiv (\mathcal{P})$. In addition, if δ is small enough, then the steps $\Delta\mu$ are not reduced by the algorithm in the* Correction step*, i.e. Newton's algorithm does not fail.*

Remark 7.4. In practice, at the end of the instruction labelled [ok], when the homotopy step has succeeded, it is possible to increase $\Delta\mu$, so that the algorithm adapts itself to the largest possible value of the homotopy step $\Delta\mu$ allowing the convergence in the Correction step.

Proof. The proof follows the ideas of [7, Prop. 8.11]. Note that the value of μ is increased only in the instruction labelled [ok] in the Update the structure step. Therefore, if the algorithm ends with $\mu = 1$, this means that all the conditions (116)–(119) are satisfied, and hence, by Lemma 6.9, θ is a vector of shooting parameters associated with a stationary point (u^1, y^1) of $(\mathcal{P}^1)$. Note that when there is no change in the structure of solutions, then Algorithm 7.1 is a classical predictor-corrector algorithm. We therefore have to show that the algorithm ends with $\mu = 1$, i.e.

- There is no failure in the Newton's algorithm in the Correction step if δ is small enough;
- The algorithm finishes off at the (finitely many by (H5)) changes in the structure of the trajectories along the homotopy path, i.e. after finitely many iterations in the Update the structure step, succeeds in finding the new structure $\mathcal{S}$ and a vector of shooting parameter $\hat{\theta}$ associated with $(u^{\bar{\mu}}, y^{\bar{\mu}})$ that satisfies the conditions (116)–(119) of Lemma 6.9.

For the current value of $\mu \in (0,1)$, assume by induction that the current value θ is a vector of shooting parameters associated with the stationary point (u^μ, y^μ) of $(\mathcal{P}^\mu)$, and that $\mathcal{S}$ denotes the corresponding structure of (u^μ, y^μ). Assume that

$$\theta \text{ and } \mathcal{S} \text{ are such that } \textit{nonreducible touch points are introduced as boundary arcs.} \tag{130}$$

(We still do not have the uniqueness of θ and $\mathcal{S}$ whenever nonessential touch points are present, that can or not be introduced in the shooting mapping.) This holds for $\mu = 0$ if p_0 is chosen sufficiently close to $p_2^0(0)$ by (H1).

Let $\bar\theta$ and $\bar\mu$ be defined as in the PREDICTION STEP. Let $\hat\theta$ be the solution of (124). By (123), $|\bar\theta - \hat\theta| \le C|\bar\mu - \mu|^2$ for some positive constant C. Since $|\bar\mu - \mu| \le \Delta\mu \le \delta$, for δ small enough, $\bar\theta$ belongs to the domain of convergence of the Newton algorithm, which converges to $\hat\theta$. Note that the constant C and the size of the domain of convergence of the Newton algorithm are uniform along the homotopy path for $\mu \in [0,1]$, see e.g. [7, Prop. 8.11], so that we do not have $\delta \to 0$.

Let us show that if δ is small enough, there is at most one passage in one of the instructions [TO→BA], [ADD TO], [REM TO], [BA→TO] before the value of μ is increased. Assume by (H5) that

$$0 \;<\; \delta \;<\; \min_{1\le k\le N-1} \tilde\mu_{k+1} - \tilde\mu_k. \tag{131}$$

If one of the tests [TO→BA], [ADD TO], [REM TO], [BA→TO] is satisfied, this means by Lemma 6.9 and (130) that the current structure $\mathcal{S}$ is not correct, and hence by (H5) and (131) there exists $k \in \{1,\dots,N\}$ such that

$$\mu \;\le\; \tilde\mu_k \;\le \bar\mu,$$

with at least one of the two above inequalities being strict, and we have $\bar\mu < \tilde\mu_{k+1}$ if $k < N$ and $\bar\mu \le 1$ if $k = N$ and $\bar\mu > \tilde\mu_N$.

Let us start by the case [TO→BA] when (125) is satisfied. This can occur only in the neighborhood of a nonreducible touch point $\bar\tau_{to}$ of $(u^{\tilde\mu_k}, y^{\tilde\mu_k})$. If $(g^{\bar\mu})^{(2)}(u^{\bar\mu}_{\mathcal{S},\hat\theta}(\tau_{to}), y^{\bar\mu}_{\mathcal{S},\hat\theta}(\tau_{to})) > 0$, a second-order Taylor expansion of $g^{\bar\mu}(y^{\bar\mu}_{\mathcal{S},\hat\theta})$ at the touch point τ_{to} shows that $g^{\bar\mu}(y^{\bar\mu}_{\mathcal{S},\hat\theta}(t)) > 0$ for t in the neighborhood of τ_{to}, $t \ne \tau_{to}$. If $(g^{\bar\mu})^{(2)}(u^{\bar\mu}_{\mathcal{S},\hat\theta}(\tau_{to}), y^{\bar\mu}_{\mathcal{S},\hat\theta}(\tau_{to})) = 0$, then τ_{to} is a nonreducible touch point. In view of (130), in both cases the structure $\mathcal{S}$ where τ_{to} is considered as a touch point is not correct. By (H6), there exist only two different neighboring structures to that of $(u^{\tilde\mu_k}, y^{\tilde\mu_k})$, so having eliminated $\mathcal{S}$, it remains only the other possible structure $\hat{\mathcal{S}}$ where $\bar\tau_{to}$ is introduced as a boundary arc. The associated new vector of shooting parameters $\bar\theta$ is obtained from θ by (126). Since we know that $\theta^{\bar\mu}$, the vector of shooting parameters associated with $(u^{\bar\mu}, y^{\bar\mu})$, is solution of

$$\mathcal{F}_{\hat{\mathcal{S}}}(\theta^{\bar\mu}, \bar\mu) = 0, \tag{132}$$

it remains to show that the Newton algorithm initialized with the value $\bar\theta$ converges to $\theta^{\bar\mu}$. Denote by $\theta^{\tilde\mu_k}_{\mathcal{S}}$ and $\theta^{\tilde\mu_k}_{\hat{\mathcal{S}}}$ the vector of shooting parameters associated with $\tilde\mu_k$ for the structures $\mathcal{S}$ and $\hat{\mathcal{S}}$, respectively, and $\bar\tau_{en}$, $\bar\tau_{ex}$, $\bar\nu^1_{\tau_{en}}$, $\bar\nu^2_{\tau_{en}}$ the shooting parameters associated with the nonreducible touch point $\bar\tau_{to}$ introduced as a boundary arc in $\theta^{\tilde\mu_k}_{\hat{\mathcal{S}}}$. Recall that the latter are given by (90)–(91). Therefore, in view of (126),

$$\begin{aligned} |\bar\theta - \theta^{\tilde\mu_k}_{\hat{\mathcal{S}}}| &\le |\theta - \theta^{\tilde\mu_k}_{\mathcal{S}}| + |\tau_{en} - \bar\tau_{en}| + |\tau_{ex} - \bar\tau_{ex}| + |\nu^1_{\tau_{en}} - \bar\nu^1_{\tau_{en}}| + |\nu^2_{\tau_{en}} - \bar\nu^2_{\tau_{en}}| \\ &\le |\theta - \theta^{\tilde\mu_k}_{\mathcal{S}}| + 2|\tau_{to} - \bar\tau_{to}| + |\nu_{\tau_{to}} - \bar\nu_{\tau_{to}}| \;\le\; 4|\theta - \theta^{\tilde\mu_k}_{\mathcal{S}}|. \end{aligned}$$

Since θ is the solution of $\mathcal{F}_{\mathcal{S}}(\theta,\mu) = 0$, it follows from Lemma 6.6 applied with $(\bar u, \bar y) = (u^{\tilde\mu_k}, y^{\tilde\mu_k})$ that there exists $\kappa > 0$ such that $|\theta - \theta^{\tilde\mu_k}_{\mathcal{S}}| \le \kappa|\mu - \tilde\mu_k| \le \kappa\delta$. By Lemma 6.6 again, there exists a constant κ' such that $|\theta^{\bar\mu} - \theta^{\tilde\mu_k}_{\hat{\mathcal{S}}}| \le \kappa'|\bar\mu - \tilde\mu_k| \le \kappa'\delta$. It follows that $|\bar\theta - \theta^{\bar\mu}| \le |\bar\theta - \theta^{\tilde\mu_k}_{\hat{\mathcal{S}}}| + |\theta^{\bar\mu} - \theta^{\tilde\mu_k}_{\hat{\mathcal{S}}}| \le (4\kappa + \kappa')\delta$. Therefore, for δ small enough, $\bar\theta$ belong to the domain of convergence of the Newton algorithm which converges to $\hat\theta := \theta^{\bar\mu}$, and all the conditions (116)–(119) are satisfied, as well as (130), so we may set $\theta := \hat\theta$, $\mu := \bar\mu$, and $\hat{\mathcal{S}} = S$ and the induction step is completed. (Here again, the constants κ, κ' can be chosen uniform w.r.t. μ along the homotopy path so that $\delta \not\to 0$.)

For the other cases, the discussion is similar so will be less detailed. In the case [ADD TO], the state constraint is violated. But then $(g^{\bar{\mu}})^{(2)}(u^{\bar{\mu}}_{\mathcal{S},\hat{\theta}}(\tau_{to}), y^{\bar{\mu}}_{\mathcal{S},\hat{\theta}}(\tau_{to})) < 0$ for all touch points τ_{to}, since otherwise we would have been in the previous case [TO→BA]. Therefore, by a second-order Taylor expansion of $g^{\bar{\mu}}(y^{\bar{\mu}}_{\mathcal{S},\hat{\theta}})$, the state constraint is not violated in the neighborhood of a touch point. Consequently, it may only be violated in the neighborhood of a nonessential touch point $\bar{\tau}_{to}$ of $(u^{\tilde{\mu}_k}, y^{\tilde{\mu}_k})$ which is not introduced in the shooting mapping. By (H6), the only other possible structure $\hat{\mathcal{S}}$ is when $\bar{\tau}_{to}$ is introduced as a touch point in the shooting mapping.

In the case [REM TO], a jump parameter associated with a touch point is negative. This cannot happen in the neighborhood of a nonreducible touch point of $(u^{\tilde{\mu}_k}, y^{\tilde{\mu}_k})$, since nonreducible touch points are assumed to be essential by (A6)(i) and Lemma 4.3. Therefore, this can only happen in the neighborhood of a nonessential touch point $\bar{\tau}_{to}$ of $(u^{\tilde{\mu}_k}, y^{\tilde{\mu}_k})$. By (H6), the only other possible structure $\hat{\mathcal{S}}$ is to remove this touch point from the shooting mapping. Finally, in the last case [BA→TO], we have a boundary arc whose entry point τ_{en} is greater than the corresponding exit point τ_{ex}. This can only happen in the neighborhood of a nonreducible touch point $\bar{\tau}_{to}$ of $(u^{\tilde{\mu}_k}, y^{\tilde{\mu}_k})$ that was converted in a boundary arc, and therefore by (H6) the only other possible structure $\hat{\mathcal{S}}$ is to introduce this nonreducible touch point as a touch point instead. We conclude with similar arguments as before that for δ small enough, the Newton algorithm initialized by $\bar{\theta}$ converges to the solution of (132), which is a vector of shooting parameters associated with the stationary point $(u^{\bar{\mu}}, y^{\bar{\mu}})$ of $(\mathcal{P}^{\bar{\mu}})$. This completes the induction step.

This shows that if δ is small enough, the algorithm follows the homotopy path, the Newton algorithm does not fail, and the algorithm ends with $\mu = 1$. By (H4), the second-order sufficient condition (65) holds and therefore (u^1, y^1) is a local solution of $(\mathcal{P})$. □

8 Remarks

Remark 8.1. It would of course be interesting to test the homotopy algorithm on numerical applications. This will be the subject of a forthcoming paper. The homotopy algorithm is based on the strong assumptions (A5) and (A6)(ii), that would have to be checked in practice in order to guarantee the validity of the algorithm, as well as the second-order sufficient condition (65). Moreover, the same restrictions as for first-order state constraints hold, see [7, Remarks 8.12 and 8.13]. In particular, a value of δ that guarantee the convergence by Prop. 7.3 is not known in practice, and may be small if the problem is ill-conditioned.

Remark 8.2. It is expected that the homotopy algorithm can be extended to vector-valued control and several state constraints of first- and second order if the constraints are linearly independent (see [22, 6]). The difficulty in the theoretical justification of the algorithm is the extension of Theorems 3.1, 4.4, and [7, Th. 2.1]. For control constraints, the extension of this homotopy algorithm is not immediate (see [7, Remark 6.3]) and is an interesting open question. In contrast, it seems not to be possible to extend this algorithm to state constraints of order greater than or equal to three, since in that case optimal trajectories typically exhibit infinitely many touch points near entry/exit of boundary arcs, see [24].

Remark 8.3. The sufficient second-order condition (65) used in Th. 5.1 is not the weakest possible since it does not take into account the curvature of the constraint. The curvature term of the constraint (see [5]) is the term with the sum in (93). It would therefore be interesting to see if Th. 5.1 is still true under the weaker second-order sufficient condition

$$\mathcal{Q}(v) - \sum_{\tau \in \bar{\mathcal{T}}^{ess}_{red}} [\bar{\eta}(\tau)] \frac{(g_y^{(1)}(\bar{y}(\tau)) z_v(\tau))^2}{g^{(2)}(\bar{u}(\tau), \bar{y}(\tau))} > 0, \quad \text{for all } v \in \mathcal{V},\ v \neq 0,\ \text{satisfying (63).} \tag{133}$$

With additional assumptions (A4)–(A5) on the structure of the trajectory and in the absence of nonreducible touch points, it was shown in [8, Th. 4.3] (see also [7]) that (133) characterizes the uniform quadratic growth condition (28), and implies L^∞-Lipchitz continuity and directional differentiability of solutions in L^r, $r < \infty$, (see [20, 8]), improving the Hölder continuity in L^∞ only obtained in Th. 5.1. Directional differentiability of all shooting parameters is also obtained. It would be interesting to extend those results in presence of nonreducible touch points as well.

Remark 8.4. Let us discuss the case when the term $\bar{\lambda}(\bar{\tau}_{to})$ defined by (58) at a nonreducible touch point $\bar{\tau}_{to}$ is positive. In that case, by Th. 4.4, a second touch point may appear for stationary points of the perturbed problem. The first idea is therefore to introduce a second touch point in the shooting mapping, and at the reference trajectory $(\bar{u}, \bar{y})$, the values of both touch points would be equal to the value of the nonreducible touch point $\bar{\tau}_{to}$. The problem is that doing so, it is easy to see that the Jacobian of the shooting mapping becomes singular (two rows are equal). Moreover, the jump parameters associated with each touch point at the reference trajectory are not well-defined, only the sum of the two jumps parameters must be equal to $\bar{\nu}_{\bar{\tau}_{to}}$. There exist indeed several zeros of the perturbed shooting function in the neighborhood of a nonreducible touch point splitting into two touch points, and one of them is such that the values of both touch points remain equal to each other (as if we had a single touch point). In that case of course the state constraint may be violated.

For this reason, it would be necessary to initialize the two touch points with distinct values and it is an open question how to do so in order to insure to be into the domain of convergence of the Newton algorithm for the new structure. To solve the academic problem in Fig. 1(b), the nonreducible touch point was first converted into a boundary arc. We thus obtained a zero of the resulting shooting function with a boundary arc satisfying $\tau_{en} < \tau_{ex}$, but the condition $\ddot{\eta}^2 \geq 0$ was of course violated. We used the obtained values of τ_{en} and τ_{ex} to initialize the two touch points, and the heuristic formula below (recall (54))

$$\nu_{\tau_{en}} := \frac{\tilde{H}_{uu}}{(g_u^{(2)})^2} g^{(3)}(\dot{u}, u, y)(\tau_{en}^-), \qquad \nu_{\tau_{ex}} := -\frac{\tilde{H}_{uu}}{(g_u^{(2)})^2} g^{(3)}(\dot{u}, u, y)(\tau_{ex}^+)$$

to initialize the associated jump parameters.

References

[1] E. L. Allgower and K. Georg. *Introduction to numerical continuation methods*, volume 45 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2003. Reprint of the 1990 edition [Springer-Verlag, Berlin].

[2] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, New York, 2000.

[3] D. Augustin and H. Maurer. Computational sensitivity analysis for state constrained optimal control problems. *Ann. Oper. Res.*, 101:75–99, 2001.

[4] P. Berkmann and H.J. Pesch. Abort landing in windshear: optimal control problem with third-order state constraint and varied switching structure. *J. of Optimization Theory and Applications*, 85, 1995.

[5] J.F. Bonnans and A. Hermant. No gap second order optimality conditions for optimal control problems with a single state constraint and control. INRIA Research Report 5837, to appear in Mathematical Programming. Online First DOI 10.1007/s10107-007-0167-8.

[6] J.F. Bonnans and A. Hermant. Second-order analysis for optimal control problems with pure state constraints and mixed control-state constraints. INRIA Research Report 6199, to appear in Annales de l'Institut Henri Poincaré (C) Analyse Non Linéaire, doi:10.1016/j.anihpc.2007.12.002.

[7] J.F. Bonnans and A. Hermant. Stability and sensitivity analysis for optimal control problems with a first-order state constraint and application to continuation methods. ESAIM Control Optim. Calc. Var. (to appear), E-first DOI: 10.1051/cocv:2008016.

[8] J.F. Bonnans and A. Hermant. Well-posedness of the shooting algorithm for state constrained optimal control problems with a single constraint and control. *SIAM J. on Control and Optimization*, 46(4):1398–1430, 2007.

[9] B. Bonnard, L. Faubourg, and E. Trelat. Optimal control of the atmospheric arc of a space shuttle and numerical simulations with multiple-shooting method. *Mathematical Models & Methods in Applied Sciences*, 15(1):109–140, 2005.

[10] A.E. Bryson, W.F. Denham, and S.E. Dreyfus. Optimal programming problems with inequality constraints I: necessary conditions for extremal solutions. *AIAA Journal*, 1:2544–2550, 1963.

[11] R. Bulirsch, F. Montrone, and H. J. Pesch. Abort landing in the presence of windshear as a minimax optimal control problem. II. Multiple shooting and homotopy. *J. Optim. Theory Appl.*, 70(2):223–254, 1991.

[12] A.L. Dontchev and W.W. Hager. Lipschitzian stability for state constrained nonlinear optimal control. *SIAM J. on Control and Optimization*, 36(2):698–718 (electronic), 1998.

[13] A.L. Dontchev and W.W. Hager. The Euler approximation in state constrained optimal control. *Mathematics of Computation*, 70:173–203, 2001.

[14] J. Gergaud and T. Haberkorn. Homotopy method for minimum consumption orbit transfer problem. *ESAIM Control Optim. Calc. Var.*, 12(2):294–310 (electronic), 2006.

[15] W.W. Hager. Lipschitz continuity for constrained processes. *SIAM J. Control Optim.*, 17(3):321–338, 1979.

[16] A. Hermant. Stability analysis of optimal control problems with a second-order state constraint. INRIA Research Report, 2007. http://hal.inria.fr/inria-00186968/fr/.

[17] K. Malanowski. Stability and sensitivity of solutions to nonlinear optimal control problems. *Journal of Applied Mathematics & Optimization*, 32:111–141, 1995.

[18] K. Malanowski. Sufficient optimality conditions in stability analysis for state-constrained optimal control. *Applied Mathematics and Optimization*, 55(2):255–271, 2007.

[19] K. Malanowski and H. Maurer. Sensitivity analysis for state constrained optimal control problems. *Discrete and Continuous Dynamical Systems*, 4:241–272, 1998.

[20] K. Malanowski and H. Maurer. Sensitivity analysis for optimal control problems subject to higher order state constraints. *Annals of Operations Research*, 101:43–73, 2001. Optimization with data perturbations, II.

[21] P. Martinon and J. Gergaud. An application of PL continuation methods to singular arcs problems. In *Recent advances in optimization*, volume 563 of *Lecture Notes in Econom. and Math. Systems*, pages 163–186. Springer, Berlin, 2006.

[22] H. Maurer. On the minimum principle for optimal control problems with state constraints. Schriftenreihe des Rechenzentrum 41, Universität Münster, 1979.

[23] H. J. Pesch. A practical guide to the solution of real-life optimal control problems. *Control and Cybernetics*, 23:7–60, 1994.

[24] H.M. Robbins. Junction phenomena for optimal control with state-variable inequality constraints of third order. *J. of Optimization Theory and Applications*, 31:85–99, 1980.

[25] S.M. Robinson. Strongly regular generalized equations. *Mathematics of Operations Research*, 5:43–62, 1980.

[26] J. Stoer and R. Bulirsch. *Introduction to Numerical Analysis*. Springer-Verlag, New-York, 1993.

Contents

Centre de recherche INRIA Saclay – Île-de-France
Parc Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 Orsay Cedex (France)

Centre de recherche INRIA Bordeaux – Sud Ouest : Domaine Universitaire - 351, cours de la Libération - 33405 Talence Cedex
Centre de recherche INRIA Grenoble – Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier
Centre de recherche INRIA Lille – Nord Europe : Parc Scientifique de la Haute Borne - 40, avenue Halley - 59650 Villeneuve d'Ascq
Centre de recherche INRIA Nancy – Grand Est : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex
Centre de recherche INRIA Paris – Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex
Centre de recherche INRIA Rennes – Bretagne Atlantique : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex
Centre de recherche INRIA Sophia Antipolis – Méditerranée : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399